From surface chlorophyll a to phytoplankton community composition in oceanic waters

Julia Uitz[1], Hervé Claustre[1], André Morel[1] and Stanford B. Hooker[2]

[1]Laboratoire d'Océanographie de Villefranche, B.P. 08, Villefranche-sur-Mer, 06 238, France

[2]NASA / Goddard Space Flight Center, Laboratory for Hydrospheric Processes, Greenbelt, MD, USA

Running title: Surface chlorophyll and algal community vertical distribution

Key words

Phytoplankton community; pigment composition; world ocean; statistical analysis; parameterization; ocean color remote sensing.

Index terms

1635 Global Change: Oceans; 1640 Global Change: Remote sensing; 4855 Oceanography: Biological and Chemical: Plankton; 4815 Oceanography: Biological and Chemical: Ecosystems, structure and dynamics; 4847 Oceanography: Biological and Chemical: Optics.

Abstract

The objective of the present study is to examine the potential of using the near-surface total chlorophyll *a* concentration (C_{surf}), as it can be derived from ocean color observation, to infer the column-integrated and the vertical distribution of the phytoplanktonic biomass, both in a quantitative way and in a qualitative way (*i.e.*, in terms of community structure). Within this context, a large HPLC (High Performance Liquid Chromatography) pigment database has been analyzed. It includes 2419 vertical pigment profiles, all sampled in Case-1 waters with various trophic states. The relationships between C_{surf} and the total chlorophyll *a* vertical distribution, as previously derived by *Morel and Berthon* [1989], are fully confirmed, as the present results coincide with the previous ones. This agreement allows to go further, namely to examine the possibility of extracting relationships between C_{surf} and the vertical composition of the algal assemblages. Thanks to the detailed pigment composition available from HPLC measurements, the contribution of three size classes (micro-, nano-, and pico-phytoplankton) to the local total chlorophyll *a* concentration can be assessed. Corroborating previous findings (*e.g.*, large species dominate in eutrophic environments, whereas tiny phytoplankton prevail in oligotrophic zones), the results lead to a statistically based parameterization. The predictive skill of this parameterization is successfully tested on a separate data set. With such a tool, the vertical total chlorophyll *a* profiles associated with each size class can be inferred from the sole knowledge of C_{surf}. By combining this tool with satellite ocean color data, it becomes conceivable to quantify on a global scale the phytoplankton biomass associated with each of the three size classes.

1. Introduction

A permanent and global monitoring of the world ocean, and in particular of the algal content of its upper layer, can only be achieved by satellite observation of the "ocean color". Ocean color, as remotely detected, provides the chlorophyll concentration restricted to the upper layer only [*Gordon and Mccluney*, 1975]. Within the context of ecological studies dealing with the vertical distribution of algal species, as well as for bio-geochemical applications involving primary production, such satellite information about the upper layer is insufficient. Indeed, the assessment of the algal biomass must be extended downward, in order to encompass the entire column where algae can live and grow. For instance, the transformation of "chlorophyll maps" as obtained from spaceborne sensors into "primary production maps", through the use of a light-photosynthesis model, requires at least that the column-integrated vegetal biomass is known, and better, that the biomass vertical distribution within the decreasingly illuminated layers can be in some way described. The terms "known" or "described" actually mean "assumed" or "predicted" with a sufficient degree of confidence, and for each "pixel" of a satellite image.

The need for extending the upper layer chlorophyll concentration was part of the motivation of the first statistical study presented by *Morel and Berthon* [1989, hereafter denoted MB89], which dealt with the shape of the phytoplankton vertical distribution and its possible relationship with the near-surface chlorophyll *a* concentration. This previous study was based on the examination of about 4000 vertical profiles of chlorophyll *a* in Case-1 waters only, and the main conclusions of this analysis were as follows:

i) The integrated biomass over the euphotic layer, and the surface concentration, are highly correlated, but in a nonlinear fashion.

ii) The vertical profiles in stratified waters can be sorted according to the surface chlorophyll *a* concentration into several trophic regimes (or trophic categories), and each of

iii) these situations exhibits a typical shape, which includes a more or less pronounced, and more or less deep, maximum; for such stratified waters, a parameterization was proposed which allows the chlorophyll *a* vertical profiles to be predicted in a continuous manner from the surface chlorophyll *a* value.

iii) And finally, well-mixed waters (*i.e.*, when the pycnocline is deeper than the euphotic depth) exhibit, as expected, substantially uniform chlorophyll *a* profiles, and have to be considered separately.

With the same general aims, the present study is firstly motivated by the considerable methodological improvement brought by the introduction, and then the systematic use, of the High Performance Liquid Chromatography (HPLC) method. Thanks to this unique technique, the determination of the chlorophyllous pigment concentration is more accurate than with the previous, sometimes diverging, fluorometric and spectrophotometric techniques, as faced with in MB89 study. The first and logical goal is to check whether the global description given in MB89 remains valid (confirmed or modified) when another data set, based on another analytical method and a more comprehensive sampling of the world ocean, is considered.

Beside the point of revisiting the previous parameterizations, which were only expressed with respect to the chlorophyll *a*, it becomes increasingly obvious that the use of this single pigment as a proxy of the algal biomass remains insufficient as far as oceanic biogeochemical cycles are to be studied and ultimately modeled. In effect, the "quality" of the phytoplankton population (namely its taxonomic composition) impacts on, or reciprocally, is a signature of, specific biogeochemical processes. For example, tiny phytoplankton are preferentially associated with the presence of regenerated forms of the nutrients they are able to utilize, whereas large phytoplankters (diatoms), which are more involved in so-called new production, develop preferentially when fresh nutrients become

available [*Eppley and Peterson*, 1979; *Malone*, 1980; *Goldman*, 1993]. A relationship between the trophic status of an oceanic system, as revealed by its near-surface chlorophyll *a* content, and the taxonomic composition of the whole algal assemblage can thus be reasonably expected [*Claustre*, 1994]. It is, therefore, timely to investigate whether a meaningful indication regarding the phytoplanktonic community structure can be inferred from the sole knowledge of the chlorophyll *a* concentration within the near-surface layer. The analysis of an HPLC pigment database offers an invaluable possibility in this direction.

As mentioned earlier, the HPLC method has brought an improved assessment of the total chlorophyll *a* concentration —for the sake of simplicity, the complex vocable "total chlorophyll *a* concentration" will be noted C; see Table 1 for a list of symbols—. In addition, and perhaps more importantly, this method allows a suite of accessory pigments (carotenoids and chlorophylls) to be determined. Many of these pigments are specific of individual phytoplanktonic taxa or groups [*Jeffrey and Vesk*, 1997]. They can thus be used as biomarkers, and eventually assigned to different phytoplankton size classes, such as micro, nano- and pico-phytoplankton [*Vidussi et al.*, 2001]. Furthermore, the vertically integrated biomasses for each algal class, or their vertical profiles, can now be determined *via* HPLC analysis, so that their respective contributions to the total standing algal stock can be assessed. The second aim of the present study is, therefore, to examine, from the analysis of HPLC data, whether some generic properties regarding the composition and vertical distribution of the phytoplankton assemblages may be inferred from the sole total chlorophyll *a* concentration in the near-surface layer, C_{surf}.

To achieve these objectives, a large HPLC pigment database has been analyzed. It encompasses stations sampled in various trophic and hydrological conditions encountered in the open ocean, and all in Case-1 waters. The data include vertical profiles of total chlorophyll *a* as well as of accessory pigments. A statistical analysis of this database is

performed to extract, and parameterize the relationships between the trophic status, as depicted by C_{surf}, and (i) the total chlorophyll *a* vertical distribution, as done in MB89, and (ii) the phytoplanktonic community composition and vertical distribution.

2. Data

Two independent pigment databases are used in the present study. Both comprise data from a large diversity of cruises to the open ocean exclusively. The first one includes 2419 vertical profiles, and the second one includes 4238 samples restricted to near-surface waters only. The former is used to perform the statistical analysis. The latter is exclusively used for validating the statistical relationships between C_{surf} and the phytoplankton composition within the surface layer. The organization and use of the databases are presented in Figure 1.

2.1. Data sources

The geographic distribution of the stations included in the database #1 is uneven (Figure 2). The Mediterranean Sea and the Atlantic and Pacific oceans are well represented (especially the subtropical gyres of the northern basins), but the South Pacific Subtropical Gyre and the Indian Ocean are inadequately represented. It must be noted that there is no data for the Arctic Ocean, and most of the stations in the Southern Ocean actually are located in the Ross Sea.

The geolocations for the second database are principally in the Atlantic Ocean. The data were collected either during cruises not belonging to the first database, or, if during the same cruises, at locations differing from those having provided the vertical profiles data (*e.g.*, surface samples taken along transects between stations).

A summary of the data collection activities for both databases is presented in Appendix A (see electronic supplement).

2.2. Overall quality control

As these samples were collected by several teams, and were analyzed in different laboratories by using a variety of instruments and HPLC methods, a quality control is required in view of coherently homogenize these data sets.

The first step of the quality control applies to all individual samples, those from the surface and those belonging to vertical profiles. The detection limit of the HPLC method depends on the sensitivity of the equipment and on the filtered volumes. Samples where total chlorophyll *a* concentration, C, is below 0.001 mg m^{-3} are rejected; for accessory pigments concentrations (denoted P for any considered pigment) values below 0.001 mg m^{-3} are reset to zero. This rejection has no real impact on the amount of significant data, because such very low pigment concentrations are normally encountered at great depths (*i.e.*, beyond the depths under investigation).

As shown by *Trees et al.* [2000], C and the sum of the concentrations of major accessory pigments are tightly correlated, and covary in a quasi-linear manner. The same kind of covariation is thus expected, and actually does exist, in the two databases used here. It can be used as a tool for a statistical test to identify and eliminate deviant data. In this study, the rejection rate is purposefully not severe, as it leads to eliminate the outliers (approximately 13%) for which the deviation exceeds three standard deviations with respect to the mean. Indeed, a balance must be maintained between the need to remove unwanted experimental errors, and the need to maintain enough data to respect the likely natural variability of the database.

The second step of the quality control deals only with the pigment vertical profiles, which are checked one by one. The systematic criteria adopted when selecting the profiles are as follows: (i) only the profiles where the euphotic depth, Z_{eu}, was reached are kept; (ii) the

uppermost sample of the profile must have been collected between the surface and 10 m depth; and (iii) each profile must include a minimal number of samples, allowing a proper description of the shape of the vertical profiles. Actually a visual inspection of each profile is necessary to eliminate those imperfectly described because of a misconceived sampling strategy or failures in the experiment. In total, this quality control process has led to the rejection of about 18% of the initially available profiles.

2.3. From accessory pigment concentration to phytoplankton size classes

The HPLC pigment analysis allows the determination of many compounds (up to 15 different pigments); some of them are redundant in term of their chemotaxinomic significance. Thus, it is appropriate to condense the information contained in these multivariable data. Along the lines of previous studies [*Claustre*, 1994; *Vidussi et al.*, 2001], "pigment indices" or pigment-based identifiers are constructed with the objective of extracting the quantitative information (concentration) and the qualitative information (taxonomic composition) encapsulated within the full suite of pigments. For this purpose, seven pigments are selected because they provide efficient discriminating tools to identify particular phytoplanktonic classes among the diversity of the oceanic algal community. These seven pigments are: fucoxanthin, peridinin, 19′-hexanoyloxyfucoxanthin, 19′-butanoyloxyfucoxanthin, alloxanthin, chlorophyll *b* and divinyl-chlorophyll *b*, and zeaxanthin. A list of these seven pigments, along with their taxonomic significance and their relationship with algal size are summarized in Table 2.

A multiple regression is performed between the total chlorophyll *a* content integrated over the euphotic layer, $\langle C \rangle_{Zeu}$, and the content of each of the seven pigments integrated over the same layer, $\langle P \rangle_{Zeu}$. This regression is highly significant ($r^2 = 0.76$; $n = 2419$; $p < 0.001$). The coefficients of this regression (Table 2) are estimates of the average ratio of total

chlorophyll *a* to each accessory pigment [see also *Gieskes et al.*, 1988; *Everitt et al.*, 1990; *Bustillos-Guzman et al.*, 1995]. These slopes are used to form the weighted sum of the diagnostic pigments concentrations, ΣDP, expressed as:

$$\Sigma DP = 1.41\ Fuco + 1.41\ Peri + 1.27\ 19'\text{-}HF + 0.35\ 19'\text{-}BF + 0.60\ Allo + 1.01\ TChl\textit{b} + 0.86\ Zea \quad (1)$$

Note that ΣDP corresponds to the C value which can be reconstructed from the knowledge of the concentration of the seven other pigments. Now, the relative proportion (%) of the biomasses associated to the three phytoplankton size classes are subsequently derived according to

$$[\%micro] = 100\ (1.41\ Fuco + 1.41\ Peri) / \Sigma DP \quad (2a)$$

$$[\%nano] = 100\ (1.27\ 19'\text{-}HF + 0.35\ 19'\text{-}BF + 0.60\ Allo) / \Sigma DP \quad (2b)$$

$$[\%pico] = 100\ (1.01\ TChl\textit{b} + 0.86\ Zea) / \Sigma DP \quad (2c)$$

These three exclusive groupings represent the aforementioned pigments indices. It is straightforward to verify that their sum is equal to 100%. The total chlorophyll *a* concentration associated with each of the three size classes is then derived according to:

$$micro\text{-}C = C\ [\%micro] \quad (3a)$$

$$nano\text{-}C = C\ [\%nano] \quad (3b)$$

$$pico\text{-}C = C\ [\%pico] \quad (3c)$$

A recent HPLC intercomparison exercise of pigment determination on natural samples, involving four laboratories, showed that C can be determined with an accuracy to within 8% [*Claustre et al.*, 2004]. Furthermore it was shown that accessory pigment ratios, as those in equations (2), are determined with a better accuracy than individual accessory pigment concentrations (because the normalization processes cancel part of the analytical bias). Thus it is expected that the use of equations (3) is very likely the most accurate approach when dealing with different data sources.

Compared to the previous parameterization proposed by *Vidussi et al.* [2001], the present one is improved in that it explicitly accounts for the differing total chlorophyll *a* to accessory pigment ratios, typical of each major phytoplankton group (all set to unity in *Vidussi et al.* [2001]).

2.4. Using and sorting the data

All of the 2419 C vertical profiles included in the database #1 (see flowchart in Figure 2) are first used to study the relationship between C_{surf} and $\langle C \rangle_{Zeu}$. In a second step, these profiles are also used to develop a parameterization of the shape of the C profiles according to the trophic status, as depicted by C_{surf}. This phase of the work tests the MB89 results against an entirely different database. The stratified waters where C_{surf} is greater than 4 mg m^{-3} are incorporated into the mixed waters category. With such concentrations, indeed, Z_{eu} is < 20 m, leading to rather uniform vertical distributions, similar to those found in mixed regime (see later).

In a second phase, the same database #1 is split into two subsets using a random process. The first subset comprises 80% of the vertical profiles (1936 stations), to be statistically analyzed for deriving the vertical average profiles of fractional C (micro-C, nano-C, and pico-C), and proposing a parameterization of these profiles. The remaining 20% (483

stations) are kept separately to validate the parameterizations. Actually, a validation is needed because such a parameterization of the fractional C profiles has never been attempted, and the validation requires a database which includes the full suite of pigments.

3. Methods

3.1. Analysis of the total chlorophyll a vertical profiles: methodological background

The present statistical analysis follows the scheme described in MB89. This method was developed for the sum of chlorophyll *a* plus phaeopigments *a* (denoted "C" in MB89), and will now be used for total chlorophyll *a* in this study, (also denoted C, as mentioned earlier). The same method will be extended without any change to any other pigment or group of pigments. Very low concentrations of phaeopigment are usually detected in oceanic waters, when determined by HPLC, and can actually be neglected.

The procedure firstly requires that the stations sampled in well-mixed waters and in stratified waters be separately considered. This sorting is made according to the ratio between the depth of the euphotic layer, Z_{eu}, and the depth of the mixed layer, Z_m, which is taken from the World Ocean Atlas 1994 [*Monterey and Levitus*, 1997]. When Z_{eu}/Z_m is ≤ 1, the sample involved is assumed to be from a well-mixed water column, and when Z_{eu}/Z_m is > 1, from a stratified water column. Z_{eu} was not usually available as an independent measurement, so it was inferred from the C vertical profile itself, using a bio-optical model. In MB89, the model of *Morel* [1988] was used. This model has been recently revised [*Morel and Maritorena,* 2001], yielding only minor changes (actually slightly larger estimates) to the derived Z_{eu} values in oligotrophic waters, and insignificant change for other waters. This new model is used in the present study.

The statistical analysis is then performed on the basis of two dimensionless quantities. A dimensionless depth, ζ, is introduced, which is obtained by dividing the actual (geometrical) depth, z, by Z_{eu}:

$$\zeta = z / Z_{eu} \tag{4}$$

Even when scaled with respect to ζ, the various profiles with differing pigment content cannot be straightforwardly compared. Another normalization is thus needed, which consists of considering dimensionless concentrations, $c(\zeta)$. These concentrations are obtained by dividing $C(\zeta)$, the local (in terms of ζ) concentration, by $\overline{C}_{Zeu}$, the average concentration within the euphotic layer, which is obtained as

$$\overline{C}_{Zeu} = (Z_{eu})^{-1} \int_0^{Zeu} C(z)\, dz, \tag{5}$$

and

$$c(\zeta) = C(\zeta) / \overline{C}_{Zeu}. \tag{6}$$

After the aforementioned double normalization has been applied, the dimensionless $c(\zeta)$ profiles can be pooled together, and compared regardless of their absolute magnitudes.

In anticipation of further statistical analyses of the profiles and their shapes, it is necessary to make the data easy to handle, *i.e.*, with regularly spaced $c(\zeta)$ values along the dimensionless depth. Consequently, for each station, a linear interpolation (with a constant increment in ζ, set to 0.1) is performed between the data of the actual profiles. This interpolation is extended down to twice Z_{eu} ($\zeta = 2$), whenever possible. The rationale for this

extension lies in the frequent occurrence of deep chlorophyll maxima, often located in oligotrophic waters in the vicinity of, or slightly beyond, Z_{eu} [*Letelier et al.*, 2004].

The dimensionless c(ζ) profiles are then sorted according to their absolute C_{surf} value into different "trophic categories", defined by successive intervals within the C_{surf} continuum. For each category, an average c(ζ) profile is computed, with its standard deviation (see Appendix B in electronic supplement).

Note that when dealing with any other pigments, or groups of pigments (like pico-C, nano-C or micro-C) the same kind of normalization can be performed: the particular local pigment (or group of pigments) concentration, P(ζ), is again normalized with respect to $\overline{C}_{Zeu}$, according to

$$p(\zeta) = P(\zeta) / \overline{C}_{Zeu} \qquad (6')$$

so that p(ζ) represents a relative proportion of a given pigment with respect to the average concentration of total chlorophyll *a* within the euphotic zone.

3.2. Total chlorophyll a vertical profiles parameterization

The regular evolution of the c(ζ) profiles as a function of C_{surf} (discussed below) allows its parameterization. In MB89, each average c(ζ) profile was modeled by using the generalized Gaussian profile [*Lewis et al.*, 1983], superimposed to a background concentration; thus, a profile was described by four parameters (three for the Gaussian profile and one for the background). A slightly modified version of this model is used here, with the objective of improving the representation of both the surface and the deepest values, the latter being generally lower than the former. This modification consists of replacing the constant

background with a linear function that decreases downward from a surface value (C_b). This substitution requires the introduction of a fifth parameter to fix the slope (s) of this decrease:

$$c(\zeta) = C_b - s\,\zeta + C_{max} \exp\left\{-[(\zeta - \zeta_{max}) / \Delta\zeta]^2\right\}, \qquad (7)$$

where C_{max} represents the maximum concentration, ζ_{max} is the depth at which this maximum occurs, and $\Delta\zeta$ depicts the width of the peak. The use of equation (7) will is explained below (Section 4.1.2.).

4. Results and Discussion

4.1. Total chlorophyll *a* biomass

4.1.1. Column-integrated total chlorophyll *a*

In MB89, an initial distinction was made between C_{pd}, the arithmetic mean concentration over the first penetration depth [*i.e.* Z_{eu}/4.6, *Gordon and Mccluney*, 1975] and C_{sat}, the weighted concentration within the same layer. This distinction has been shown (in MB89) to be useless, as these quantities are in effect confounded. They are replaced here by a unique quantity (denoted C_{surf}, as mentioned earlier) which represents the arithmetic average over the first penetration depth. The column total chlorophyll *a* content integrated over the euphotic zone, $\langle C \rangle_{Zeu}$, is then studied as a function of C_{surf}. The log-log plots of these quantities (Figure 3) suggest nonlinear relationships of the form

$$\langle C \rangle_{Zeu} = A\,(C_{surf})^B \qquad (8)$$

are appropriate (as in MB89), and that the exponent B is notably below unity.

In the plot for stratified regimes (Figure 3a), a distinct change in the slope seems to occur around $C_{surf} \approx 1$ mg m^{-3}, so two separate regression analyses on each side of this threshold are performed (Table 3). Note that a similar discontinuity was already noted in MB89. In the domain $C_{surf} < 1$ mg m^{-3}, the scatter of the points is more important than in the $C_{surf} > 1$ mg m^{-3} domain, and reflects the variability associated with the existence of a deep chlorophyll maximum typical of oligotrophic conditions.

In mixed waters, the log-log plot of $\langle C \rangle_{Zeu}$ *vs* C_{surf} (Figure 3b) shows a unique linear trend, a steeper slope, and a low scatter in the data, as a result of the rather homogenous pigment distribution. The corresponding power law is highly significant (Table 3).

The present results are very consistent with, and independently confirm those of MB89 (the dashed lines in Figure 3). The exponent in the relationship for stratified waters and when $C_{surf} \leq 1$ mg.m^{-3} is slightly lower than in the MB89 relationship (0.357 *vs* 0.425), which leads, in oligotrophic waters, to higher $\langle C \rangle_{Zeu}$ values than in the MB89 study. The differences are statistically insignificant. For mixed waters, the present relationship is essentially identical to that of stratified waters with $C_{surf} > 1$ mg m^{-3}. The closeness of the exponents for stratified waters when $C_{surf} > 1$ mg m^{-3} and for mixed waters (0.615 *vs* 0.538 respectively) was also reported in MB89. It emphasizes that (notwithstanding the deep stratification) the algal biomass in eutrophic conditions remains homogeneously distributed within the euphotic layer, essentially as a consequence of its reduced thickness.

4.1.2. Total chlorophyll *a* vertical profiles

On the basis of the C_{surf} values, the vertical C profiles have been sorted as in MB89, except that more trophic categories, namely nine for the stratified waters (S1 to S9), and five for the mixed waters (M1 to M5) have been presently considered, instead of seven and two in MB89, respectively. The C_{surf} intervals selected for defining each category are provided in

Table 4; note also that the lower limits are very different, with < 0.04 mg m^{-3} for S1, and < 0.4 mg m^{-3} for M1. This more detailed categorization scheme was made possible because the data used in this study are rather uniformly distributed, which permitted a significant number of profiles to be enlisted in each of the nine or five trophic categories.

Stratified waters

In the stratified regime, the shapes of the average dimensionless profiles exhibit a remarkably regular evolution along the trophic status (Figure 4a). When C_{surf} increases, the change in shape is characterized by i) a smooth ascent of the C maximum, from a depth corresponding to about Z_{eu} for S1, up to the near-surface for S9; ii) a regular decrease in its magnitude; and iii) a widening and weakening of the maximum that tends to vanish in eutrophic waters. It is worth noting that the regular changes of the average profiles are similar to those described in MB89 (their categories a to g, for stratified regimes).

The average dimensionless profiles are used in conjunction with equation (7). The fitting procedure (a Newton-type algorithm from the library of the R software [*R Development Core Team*, 2004]) allows the five parameters involved in equation (7) to be derived for each category (for the corresponding average C_{surf} value), from S1 to S9 (Table 4). The modeled profiles (Figure 4b) compare well with the average profiles. The values of the five parameters for each of the nine trophic categories are tabulated in Table 5.

The regular evolution in shape, as described above, suggests the parameterization can be used in a continuous fashion. For any C_{surf} value, the values of the five parameters can be interpolated between the discrete values specific of each stratified waters category. This interpolation can be computed linearly (present study) or by using fitted polynomials (as done in MB 89).

The dimensionless profiles can be easily restored to their dimensional concentrations and depths, by using the actual Z_{eu} and $\overline{C}_{Zeu}$ values. Both are retrieved using the relationships of equation (8), linking $\langle C \rangle_{Zeu}$ to C_{surf}, and Z_{eu} to $\langle C \rangle_{Zeu}$ [equation (6) in *Morel and Maritorena*, 2001].

These restored (or "rescaled") profiles (Figures 4c, d) provide a direct visual appraisal of the evolution in the true profiles when the surface concentration changes. For instance, with a C_{surf} value increasing from about 0.03 to 3.0 mg m^{-3} (S1 to S9), the chlorophyll maximum increases from 0.17 to 3 mg m^{-3}, whereas Z_{eu} decreases from 120 m to 26 m.

Mixed waters

In the mixed regime, the average normalized C profiles (Figure 5a) show a rather featureless shape: no maximum (except for M1), a constant concentration which remains approximately equal (±10%) to the average concentration within the euphotic layer, and finally a systematic decrease of the average concentrations (by about a factor of 2) between Z_{eu} and 2 Z_{eu}. The values observed at 2 Z_{eu}, however, are still about 40% of the average value within the euphotic column, instead of 20% in stratified waters. The rescaled profiles (Figures 5b and c) show that C varies from approximately 0.15 to about 6-7 mg m^{-3}. The first situation (M1) is encountered during the winter vertical mixing in the North Atlantic, and also in the Mediterranean Sea, where the mixed layer depth can reach 200 m [with $C \approx 0.15$ mg m^{-3}; *Marty et al.*, 2002]. The category M5 is typical of upwelling conditions, with a mixed layer extending deeper than the thin euphotic layer, and C ranging from 4 to about 30 mg m^{-3}. A remarkable agreement between the present study results and those obtained in MB89 is noticed for these mixed regimes.

4.1.3. Closer comparison with the MB89 parameterization

For such a comparison, forty C_{surf} values, regularly spaced in terms of their logarithms, have been generated within the interval 0.04 and 3 mg m^{-3}. For each of these C_{surf} values, two C profiles are reconstructed by using the MB89 parameterization and the present one. Then, three main characteristics of these profiles are compared: the average concentration, $\langle C \rangle_{Zeu}$, the peak value, C_{max}, and the actual depth of this peak, Z_{max} (Figure 6). Globally, both parameterizations generate similar values for these shape characteristics. The only significant divergence is for Z_{max} which, in the middle range (20-30m), is smaller with the present parameterization than it was by using MB89. This is partly a result of the differing bio-optical models used to infer Z_{eu}. For mixed waters (without a deep maximum), there is no equivalent comparison; the comparison actually reduces to the relationship between $\langle C \rangle_{Zeu}$ and C_{surf}, which has already been recorded as being identical in the two studies (Figure 3b).

4.2. Phytoplankton composition and vertical distribution

Only 80% of the database #1 is now used to study the phytoplankton composition and its vertical distribution (Figure 2). A required first step in this analysis is to verify that this subset represents the whole set without significant bias. With this objective in mind, the same statistical analysis for stratified and mixed waters is repeated on this partial data set and compared to the results obtained when using the whole set. Regarding the four characteristics representative of each trophic category (*i.e.*, the average C_{surf}, $\overline{C}_{Zeu}$, $\langle C \rangle_{Zeu}$, and Z_{eu}), no significant change can be detected between the subset and the entire set —the average values and the associated standard deviations for each of the above characteristics, obtained for the subset, deviate by less than 1% from those derived for the entire data set. Consequently, the following analysis based on the 80% subset can be safely undertaken.

4.2.1. Micro-, nano-, and pico-phytoplankton in stratified waters

Column-integrated contents

By using the detailed pigment composition together with equations (2) and (3), the fractional C, related to each of the three size classes, can be computed at each depth and for each profile. These profiles are vertically integrated down to the euphotic depth, and then averaged within each trophic category (S1 to S9). The comparison of these column-integrated fractional C with $\langle C \rangle_{Zeu}$ (all populations confounded) is shown as relative proportions (in percent), or as absolute values (in mg m^{-2}), in Figures 7a and 7b respectively, with both plotted as a function of C_{surf}.

In the most oligotrophic waters ($C_{surf} \approx 0.03$ mg m^{-3}), pico- and nano-phytoplankton are of equal importance; all together they amount to 90% of the total chlorophyll *a* integrated content. In contrast, micro-phytoplankton dominate (75%) in eutrophic waters ($C_{surf} \approx 3$ mg m^{-3}), whereas nano- and pico-phytoplankton do not represent more than 20% and 5% of $\langle C \rangle_{Zeu}$, respectively. In mesotrophic waters ($C_{surf} \approx 0.3$ mg m^{-3}), nano-phytoplankton are predominant (50%), while pico- and micro-plankton account for the remainder in similar proportions (25% each). When the absolute contents (mg m^{-2} in Figure 7b) are considered, a simple picture emerges; the increase in the C values along the trophic continuum (two orders of magnitude in C_{surf}) is essentially caused by the increase in the micro-phytoplankton biomass (by a factor of seven); meanwhile the pico-phytoplankton population appears as being a rather constant background (≈ 5 mg m^{-2}), whereas the nano-phytoplankton population experiences an increase (actually a tripling) from oligotrophic to eutrophic conditions.

The observed variations in the phytoplankton assemblages, here quantified as a function of the trophic gradient, are in agreement with previous microscopic observations and ecological studies. It is widely acknowledged that the phytoplankton size is associated with

particular types of trophic environment [*Malone*, 1980; *Chisholm*, 1992]; pico-phytoplankton would be the main biomass contributor in extremely oligotrophic systems (*e.g.*, subtropical gyres), whereas diatoms, which dominate the micro-phytoplankton class, are predominant in eutrophic areas (*e.g.*, upwellings). Furthermore the present quantitative analysis reveals the rather global ubiquity and stability (20-50% of the biomass) of the nano-phytoplankton biomass, which is in agreement with previous observations of 19′-hexanoyloxyfucoxanthin (the main pigment contributing to nano-C) being an ubiquist pigment in many oceanic situations [*Ondrusek et al.*, 1991; *Vidussi et al.*, 2001].

When the relative proportions are analyzed for the surface layer only (Figure 7c) and compared to those recorded for the euphotic layer (Figure 7a), noticeable differences are seen for the S2-S4 categories: the pico-phytoplankton proportion is enhanced at the expense of the nano-phytoplankton.

Vertical Profiles

In stratified waters, the average dimensionless profiles of micro-C, nano-C, and pico-C exhibit significant and regular evolutions according to the trophic regime (Figures 8a-c). The profiles relative to the pico- and nano-phytoplankton classes hold much in common with those for the total phytoplankton population (Figure 4a), and are characterized by a distinct deep maximum becoming shallower and smoother as C_{surf} increases. This maximum is slightly sharper for pico- than for nano-phytoplankton in most oligotrophic waters (S1 to S4); nevertheless, the peaks have approximately the same intensity. Note that for the most oligotrophic category (S1), the dimensionless concentrations of pico- and nano-phytoplankton are each about equal to $\overline{C}_{Zeu}$, which is possible because the corresponding total chlorophyll *a* value is about twice $\overline{C}_{Zeu}$ (cf. Figure 4a). Globally, the deep C maximum for these waters is almost entirely made up of pico- and nano-phytoplankton assemblages

(each with about 40%), in agreement with the column-integrated contents (see above); the contribution of the micro-phytoplankton is extremely low (< 20%) with rather featureless profiles.

In eutrophic conditions (S8 and S9), a micro-phytoplankton maximum tends to develop near the surface, while the pico-C and nano-C profiles become uniform, with low relative values. The nano-phytoplankton is predominant everywhere in the water column for intermediate trophic regimes, (*i.e.*, when C_{surf} is around 0.2-0.5 mg m^{-3}; see S5 and S6).

Parameterizing the micro-C, nano-C, and pico-C profiles is made as above for C, according to equation (7). The profiles reconstructed with the appropriate parameters (Table 5) are displayed in Figures 8d- f. They reproduce fairly well the profiles derived from the statistical analysis.

The same profiles, restored in their actual concentration and depth values, are displayed in Figure 9. The absolute magnitude of the pico-phytoplanktonic peaks remains about the same (0.09-0.11 mg m^{-3}) for the first four classes (S1 to S4), in contrast to the nano-phytoplanktonic peaks which steadily increase (0.08 to 0.14 mg m^{-3}).

4.2.2 Micro-, nano-, and pico-phytoplankton in well-mixed waters

Global ocean

About one third of the mixed waters stations (184 over 479) were located in the Southern Ocean (south of 60°S). These data are examined separately for reasons explicated thereafter. The 295 remaining stations from the global ocean (Southern Ocean excluded) are analyzed according to the same trophic categories as defined above (M1 to M5), with respect to their vertically integrated content, and then with respect to the shape of their profiles.

The integrated contents of micro-C, nano-C, and pico-C, expressed as percentage of the total content, are displayed in Figure 10a (to be compared with Figure 7a). The figure shows

there is no notable difference between the mixed and stratified waters regarding their respective proportions of pico-, nano-, and micro-phytoplankton contents, except a lower contribution of pico-phytoplankton in oligotrophic conditions (30% *vs* 50% in mixed and stratified waters, respectively).

As expected from the profiles of total chlorophyll *a* (Figure 5b), the individual profiles for pico-, nano-, and micro-phytoplankton (Figures 11a-c) are also rather featureless, essentially uniform within the euphotic layer, and decreasing below (ζ from 1 to 2). Regarding the size composition, they bear some resemblance with those in stratified waters. Indeed, in low C waters (C_{surf} < 0.4 mg m^{-3}), the micro-phytoplankton fraction is at its minimum (10% of $\overline{C}_{Zeu}$), while the pico- and nano-phytoplankton dominate (with a smooth peak near ζ = 0.8). The latter is relatively more abundant (50% *vs* 30%) in contrast to what happens in stratified waters. In eutrophic regimes, the micro-phytoplankton fraction largely dominates (more than 80%), and this dominance is similar to what was described for stratified systems (Figure 7a). Note that, in such high C stratified systems, the depth of the euphotic zone is reduced to such a point that waters could be equally considered as mixed ones (according to Z_{eu}/Z_m).

Particular case of the Southern Ocean

Southern (mixed) waters are treated separately in this study. A preliminary analysis of the data set showed that, contrarily to temperate latitudes where a C increase was generally associated with a Fuco (diatoms) increase, in the Southern Ocean, 19'-HF (prymnesiophytes) increase was usually observed during bloom periods. About 80% of these data originate from the Ross Sea. This area is known to sustain *Phaeocystis* blooms in relatively mixed conditions, whereas diatom blooms occur in stratified areas [*Goffart et al.*, 2000; *Boyd*, 2002; *Arrigo et al.*, 2003]. Studies conducted in other southern areas also reported that

phytoplankton blooms can be dominated either by nano-phytoplankton species (*e.g.*, *Phaeocystis spp.*, cryptophytes, and nanoflagellates) or by diatoms [*Moline and Prézelin*, 1996; *Claustre et al.*, 1997]. More generally, these observations and the subsequent partition of the data set (mixed waters) into two subsets (global and southern waters) are consistent with a recent body of literature which emphasizes the particular bio-optical status of polar waters [*Stramska et al.*, 2003; *Cota et al.*, 2003], possibly related to taxonomic composition differences [*Sathyendranath et al.*, 2001], when compared to temperate or tropical environments.

The partition of the data set between southern and temperate (*i.e.*, global without southern) waters appears *a posteriori* supported by the examination of Figure 10b (to be compared with Figure 10a). Indeed, in southern waters, the relative abundance of the three size classes is significantly different, and the pattern of the nano- and micro-phytoplankton contributions is reversed. Actually, micro-phytoplankton dominate at low C_{surf} and decrease to the profit of nano-phytoplankton when C_{surf} increases. At C_{surf} greater than 2 mg m^{-3}, nano-phytoplankton represent up to 60 to 80% of the biomass in southern waters (while diatoms would represent about 90% of the biomass in temperate waters). The pico-phytoplankton biomass, which never exceeds 10% of the total chlorophyll *a* content, is distinctly lesser in southern waters than in any other areas.

The vertical profiles in the Southern Ocean (Figures 11d-f) present a relatively uniform distribution (down to 2 Z_{eu}) for the nano-phytoplankton assemblages, and also show that this population is preponderant. A micro-phytoplankton deep maximum seems to develop above the euphotic depth, in low C_{surf} waters (M1 class). A closer examination of the corresponding profile shows a relatively high variability around the mean maximum value ($\approx 0.8\ \overline{C}_{Zeu} \pm 0.4$).

Parameterization in mixed waters

In the mixed regime, the average proportion of each phytoplankton class within the euphotic layer is dependent on the geographical location (global *vs* southern) and on the trophic status, as a function of C_{surf}. Furthermore, with the exception of a few situations, the vertical distribution of the micro-C, nano-C, and pico-C generally present weak changes with depth, at least within the euphotic layer. It is reasonable, therefore, to use the results presented in Figure 10 as a basis for inferring, from the sole knowledge of C_{surf}, the vertical distribution of micro-C, nano-C, and pico-C with depth. The fractional C is simply calculated by multiplying C_{surf} by the corresponding proportion, and this concentration is thereafter assumed to be homogenous from the surface down to Z_{eu}. The proportions for each mixed trophic category are tabulated in Table 6.

4.2.3 Validation of the method for predicting the micro-C, nano-C and pico-C profiles

Vertical distribution

The fraction of the HPLC database (20%, including 483 micro-C, nano-C and pico-C vertical profiles), which was kept for the validation process, is presently used (Figure 2). The euphotic depth is computed from the C_{surf} value, and compared to Z_m, to determine if the water column is stratified or mixed. For stratified waters, the C_{surf} value is injected into the relevant parameterization to generate the dimensionless vertical profiles which are then restored in physical units (by using the actual Z_{eu} and $\overline{C}_{Zeu}$ values). For mixed waters, vertical profiles are simply derived by extrapolating to depth the product of C_{surf} by the relative proportion of each group. By using these profiles, the column-integrated content (for micro-, nano-, and pico-phytoplankton) can be predicted (for the 0-Z_{eu} interval), and directly compared to the measured contents (Figure 12). For the micro-C and the nano-C (Figures 12a-b), the points are rather regularly distributed along the 1:1 line over a large range of

concentrations (≈ 0.5-100 mg m^{-2} for micro-phytoplankton, and ≈ 2-100 mg m^{-2} for nano-phytoplankton). The distribution is apparently less satisfying for pico-phytoplankton, probably because the biomass of this group varies over a much more restricted range (2-8 mg m^{-2} predicted *vs* 1.5-15 mg m^{-2} measured).

For stratified waters, the parameterization is further evaluated through the analysis of two descriptors of the profile shape: the maximum value of the micro-C, nano-C and pico-C profiles, and the depth Z_{max} of this maximum (Figures 12d-j). The same comments given for Figures 12a-c apply here. The agreement between measured and predicted shape parameters is satisfactory as far as the micro- and the nano-phytoplankton are concerned. The situation is less favorable for pico-phytoplankton, but this is largely because of a weaker parameterization of this class in southern waters. Were these waters disregarded, the quality of the parameterization would be nearly similar for the three groups.

Surface layer

This validation uses the separate databank (database #2, see Figure 2) which includes 4238 surface samples. The actual C concentrations (sampled between 0 and 10 m) are used as C_{surf}. The respective proportions of micro-, nano-, and pico-phytoplankton are retrieved for each C_{surf}, by using the results of Figure 7c for stratified waters, and of Figure 10 for mixed waters. The micro-C, nano-C, and pico-C are then computed by multiplying the corresponding proportions with C_{surf}, and finally compared to the measured values (Figure 13). For the global ocean (southern waters excluded), the agreement is excellent over a wide range of C_{surf} values for the micro- and nano-phytoplanktonic groups, and also for the pico-phytoplankton, despite a much restricted range of variation. Again, the parameterization is less satisfying for southern waters.

4.2.4. Example of application of the proposed parameterization

The relevance of the parameterization developed as part of this study can be shown by mapping the distribution of the three phytoplankton size classes at the global scale. A SeaWiFS monthly composite (for June 2000) is taken as an example. The C value associated with each pixel (C_{surf}) is used to infer the proportions, as well as the vertical distribution, of the biomass corresponding to the three phytoplankton size classes. A preliminary partition of the ocean into mixed and stratified waters is necessary before using the appropriate parameterization (Tables 5 and 6). The results display the typical phytoplankton signatures of the main oceanic entities (Figures 14a-c).

Pico-phytoplankton appear to be the dominant group within the subtropical gyres where they can represent up to 50% of the biomass. In high latitude environments, however, their contribution is less than 10%, especially in the northern hemisphere. The reverse situation is observed for micro-phytoplankton, which always represent less than 20% (sometimes less than 10%) in the gyres, and, in contrast, 40-70% of the biomass in the North Atlantic (latitudes greater than 40°N), where they form the dominant size class at the end of the spring bloom. In between, nano-phytoplankton represent an extremely stable biomass everywhere (30%-50% of the total).

The analysis can be further extended by considering the vertically integrated content of micro-C, nano-C and pico-C within the euphotic layer (Figures 14d-f). Micro-phytoplankton (likely diatoms) definitely present the largest range of biomass variation with values close to 0 in the center of the subtropical gyres and up to 80 mg m^{-2} in the North Atlantic Ocean. Pico-phytoplankton biomass varies within a narrow range of concentrations (7-8 mg m^{-2}). As a result of its relative stability, the nano-phytoplankton distribution typically mimics that of the total biomass.

5. Conclusions

Although the MB89 study was based on a substantial amount of data (3806 profiles), thus leading to rather robust results, it is still appropriate to test this robustness by considering an independent data set. Such a test is particularly required if an extension of the same approach is envisioned, dealing with other pigments, or groups of pigments, instead of the total chlorophyll *a* alone.

The results of this test are highly satisfactory. The relationships between C_{surf} and the total chlorophyll *a* column-integrated content in both stratified and mixed waters of this study coincide perfectly with those obtained in MB89. This comforting agreement calls for two comments. The methodological change when determining C has no discernible impact upon the results, which is an indirect consolidation of the previous measurements, which were mainly based on the fluorometric method. Such global relationships testify to the existence of robust global trends over a wide range of trophic conditions. The scattering of the points (see Figure 3), however, simultaneously demonstrates that any relationship of that sort is not as well defined when a restricted range of C is considered [*e.g.* discussion in *Siegel et al.*, 2001].

Regarding the regular and progressive evolution of the shapes of the vertical profile according to C_{surf}, the strong resemblance (actually a close numerical agreement) between the previous results and those presently obtained is very encouraging. This confirmation reinforces the notion of a continuity in the relationship between C_{surf} and the featured vertical distribution typical of stratified waters (replaced by uniform distributions in mixed waters). Such a continuity can be clearly discerned when the statistical analysis is performed by using dimensionless quantities, for both the depth and the concentration, whereas it remains obscure when the physical quantities are directly considered. For modeling purposes, including those for primary production based on satellite surface chlorophyll data, such a continuity is helpful and easy to handle from a numerical point of view.

The realization of the continuity is also supported by the evidence of regular changes within the pigment composition, the continuum of algal cell size, and the vertical layering of the dominant populations. The results dealing with the phytoplanktonic assemblages and the predominance of algal groups according to the trophic conditions, confirms existing qualitative descriptions, generally based on local observations. Though, owing to the statistical analysis, the present study extends the previous knowledge, and permits its generalization to most oceanic situations.

The proposed statistical relationships, as all statistical products, may fail on a case-by-case basis, and their ubiquity is certainly doubtful in high latitudes. Indeed, polar zones are largely underrepresented in the present data set, with the exception of the Ross Sea which has received a particular attention because of its relatively high productivity. The hydrological and climatological conditions, prevailing in southern waters, have been found to favor particular phytoplankton assemblages with respect to those growing in temperate regions. A special statistical analysis performed on a more specific data set must, therefore, be considered for these particular waters, as well as for Arctic waters not represented here.

Finally, this study constitutes a first step in the synthesis of HPLC data collected during the last decade. These data, originating from many sources, provide an initial framework to comprehensively describe the distribution of the phytoplankton communities at global scales, which has become increasingly needed to appraise how the algal biomass interacts with carbon fluxes. The generalized trends in the distribution of phytoplankton size classes according to the trophic status are not only confirmed in this study at global scales, but are also parameterized in a practical way; it thus becomes possible to propose a first assessment of the phytoplankton community composition for the world ocean from ocean color satellite observation (Figure 14). The transformation of "maps of phytoplankton size class biomass" into "maps of primary production specific to each size class", through the use

of bio-optical models, appears now feasible (to the extent that the photosynthetic parameters typical of the different phytoplankton size classes can be selected). Such maps of "specific primary production" could in turn be very useful to improve the parameterization and the validation of recent biogeochemical models which include different phytoplankton compartments [*e.g. Moore et al.*, 2002; *Aumont et al.*, 2003]).

Acknowledgements

People who kindly contributed to the implementation of the database are gratefully acknowledged: J. Aiken, K. Arrigo, R. Barlow, R. Bidigare, C. Cailliau, Y. Dandonneau, G.R. DiTullio, L. Dransfield, E. Head, H. Higgins, G. Kraay, J.C. Marty, C. Omachi, K. Oubelkheir, J. Ras, S. Roy, V. Stuart, C. Targa, D. Thibault, C. Trees, L. Van Heukelem, M. Veldhuis, F. Vidussi, G. Westbrook, S. Wright, M. Zapata. Special thanks to J. Ras for her implication in the early stages of the pigment database merging process. We also thank D. Antoine for his helpful advices in the ocean color maps process, and B. Gentili for his help in programming.

References

Arrigo, K. R., D. L. Worthen, and D. H. Robinson (2003), A coupled ocean-ecosystem model of the Ross Sea: 2. Iron regulation of phytoplankton taxonomic variability and primary production, *J. Geophys. Res., 108*(C7), 3231, doi: 10.1029/2001JC000856.

Aumont, O., E. Maier-Reimer, S. Blain, and P. Monfray (2003), An ecosystem model of the global ocean including Fe, Si, P colimitations, *Global Biogeochem. Cycles, 17*(2), 1060, doi:10.1029/2001GB001745.

Boyd, P. W. (2002), Environmental factors controlling phytoplankton processes in the Southern Ocean, *J. Phycol., 38*, 844-861.

Bustillos-Guzman, J., H. Claustre, and J. C. Marty (1995), Specific phytoplankton signatures and their relationship to hydrographic conditions in the coastal northwestern Mediterranean Sea, *Mar. Ecol. Prog. Ser., 124*(1-3), 247-258.

Chisholm, S. W. (1992), Phytoplankton size, in *Primary Productivity and Biogeochemical Cycles in the Sea*, edited by P. G. Falkowski and A. D. Woodhead, pp.213-237, Plenum Press, New York.

Claustre, H. (1994), The trophic status of various oceanic provinces as revealed by phytoplankton pigment signatures, *Limnol. Oceanogr., 39*(5), 1206-1210.

Claustre, H., S. B. Hooker, L. Van Heukelem, J. F. Berthon, R. G. Barlow, J. Ras, H. Sessions, C. Targa, C. S. Thomas, D. van der Linde, and J. C. Marty (2004), An intercomparison of HPLC phytoplankton methods using in situ samples: Application to remote sensing and database activities, *Mar. Chem., 85*(1-2), 41-61.

Claustre, H., M. A. Moline, and B. B. Prezelin (1997), Sources of variability in the column photosynthetic cross section for Antarctic coastal waters, *J. Geophys. Res., 102*(C11), 047-25.

Cota, G. F., W. G. Harrison, T. Platt, S. Sathyendranath, and V. Stuart (2003), Bio-optical properties of the Labrador Sea, *J. Geophys. Res., 108*(C7), 32,28 doi: 10.1029/2000JC000597.

Eppley, R. W., and B. J. Peterson (1979), Particulate organic matter flux and planktonic new production in the deep ocean, *Nature, 282*, 677-680.

Everitt, D. A., S. W. Wright, J. K. Volkman, D. P. Thomas, and E. J. Lindstrom (1990), Phytoplankton community compositions in the western Equatorial Pacific determined from chlorophyll and carotenoid pigment distributions, *Deep Sea Res., 37*(6A), 975-997.

Gieskes, W. W. C., G. W. Kraay, A. Nontji, D. Setiapermana, and Sutomo (1988), Monsoonal alternation of a mixed and a layered structure in the phytoplankton of the euphotic zone of the Banda Sea (Indonesia): A mathematical analysis of algal pigment fingerprints, *Neth. J. Sea Res., 22*(2), 123-137.

Goffart, A., G. Catalano, and J. H. Hecq (2000), Factors controlling the distribution of diatoms and Phaeocystis in the Ross Sea, *J. Mar. Syst., 27*(1-3), 161-175.

Goldman, J. C. (1993), Potential role of large oceanic diatoms in new primary production, *Deep Sea Res., 40*, 159-168.

Gordon, H. R., and W. R. McCluney (1975), Estimation of the Depth of Sunlight Penetration in the Sea for Remote Sensing, *Applied Optics, 14*(2), 413-416.

Jeffrey, S. W., and M. Vesk (1997), Introduction to marine phytoplankton and their pigment signatures, in *Phytoplankton pigments in oceanography*, edited by S. W. Jeffrey, R. F. C. Mantoura and S. W. Wright, pp.37-84, UNESCO, Paris.

Letelier, R. M., D. M. Karl, M. R. Abbott, and R. R. Bidigare (2004), Light driven seasonal patterns of chlorophyll and nitrate in the lower euphotic zone of the North Pacific Subtropical Gyre, *Limnol. Oceanogr., 49*(2), 508-519.

Lewis, M. R., J. J. Cullen, and T. Platt (1983), Phytoplankton and thermal structure in the upper ocean: Consequences of nonuniformity in chlorophyll profile, *J. Geophys. Res., 88*, 2565-2570.

Malone, T. C. (1980), Algal size, in *The Physiological Ecology of Phytoplankton*, edited by I. Morris, pp.433-463, University of California, Berkeley.

Marty, J. C., J. Chiaverini, M. D. Pizay, and B. Avril (2002), Seasonal and interannual dynamics of nutrients and phytoplankton pigments in the western Mediterranean Sea at the DYFAMED time-series station (1991-1999), *Deep Sea Res., 49*(11), 1965-1985.

Moline, M. A., and B. B. Prézelin (1996), Long-term monitoring and analyses of physical factors regulating variability in coastal Antarctic phytoplankton biomass, in situ productivity and taxonomic composition over subseasonal, seasonal and interannual time scales, *Mar. Ecol. Prog. Ser., 145*(1-3), 143-160.

Monterey, G., and S. Levitus (1997), *Seasonal Variability of Mixed Layer Depth for the World Ocean, NOAA Atlas NESDIS*, vol. 14, 96 pp., U.S. Gov. Printing Office, Washington, D.C.

Moore, J. K., S. C. Doney, J. A. Kleypas, D. M. Glover, and I. Y. Fung (2002), An intermediate complexity marine ecosystem model for the global domain, *Deep Sea Res., 49*(1-3), 403-462.

Morel, A. (1988), Optical modeling of the upper ocean in relation to its biogenous matter content (case 1 waters), *J. Geophys. Res., 93*(C9), 10,749-10,768.

Morel, A., and J. F. Berthon (1989), Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications, *Limnol. Oceanogr., 34*(8), 1545-1562.

Morel, A., and S. Maritorena (2001), Bio-optical properties of oceanic waters: A reappraisal, *J. Geophys. Res., 106*(C4), 7163-7180.

Ondrusek, M. E., R. R. Bidigare, S. T. Sweet, D. A. Defreitas, and J. M. Brooks (1991), Distribution of phytoplankton pigments in the North Pacific Ocean in relation to physical and optical variability, *Deep Sea Res., 38*(2), 243-266.

R Development Core Team (2004), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-00-3, URL http://www.R-project.org.

Sathyendranath, S., G. Cota, V. Stuart, H. Maass, and T. Platt (2001), Remote sensing of phytoplankton pigments: a comparison of empirical and theoretical approaches, *Int. J. Remote Sens., 22*(2 & 3), 249-273.

Siegel, D. A., T. K. Westberry, M. C. O'Brien, N. B. Nelson, A. F. Michaels, J. R. Morrison, A. Scott, E. A. Caporelli, J. C. Sorensen, S. Maritorena, S. A. Garver, E. A. Brody, J. Ubante, and M. A. Hammer (2001), Bio-optical modeling of primary production on regional scales: the Bermuda BioOptics project, *Deep Sea Res., 48*(8-9), 1865-1896.

Stramska, M., D. Stramski, R. Hapter, S. Kaczmarek, and J. Ston (2003), Bio-optical relationships and ocean color algorithms for the north polar region of the Atlantic, *J. Geophys. Res., 108*(C5), 3143, doi: 10.1029/2001JC0001195.

Trees, C. C., D. K. Clark, R. R. Bidigare, M. E. Ondrusek, and J. L. Mueller (2000), Accessory pigments versus chlorophyll *a* concentrations within the euphotic zone: A ubiquitous relationship, *Limnol. Oceanogr., 45*(5), 1130-1143.

Vidussi, F., H. Claustre, B. B. Manca, A. Luchetta, and J. C. Marty (2001), Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter, *J. Geophys. Res., 106*(C9), 19,939-19,956.

Figures captions

Figure 1: Geographic location of the 2419 stations used in the present study. The size of the black dots indicates the number of stations within a square of 10° longitude by 10° latitude. The small dot corresponds to less than 10 stations; the medium to 10-100 stations; and the large to more than 100 stations.

Figure 2: Flowchart detailing the partition and the use of the two data sets for data analysis as well as for validation of the developed parameterization.

Figure 3: Relationship between the average total chlorophyll *a* concentration within the surface layer, C_{surf}, and the total chlorophyll *a* integrated content within the euphotic layer, $\langle C \rangle_{Zeu}$, for stratified (a) and mixed (b) waters. The solid lines represent equations (8) while the dashed lines correspond to the regression lines models of MB89.

Figure 4: C vertical profiles for stratified waters. Average dimensionless C profiles for each trophic category (a), modeled dimensionless C profiles for each trophic category (b), rescaled C profiles for categories S1 to S4 (c), and rescaled C profiles for categories S5 to S9 (d). For rescaled profiles of each trophic category, the position of Z_{eu} is identified by an arrow associated with the symbol relevant to the trophic category.

Figure 5: C vertical profiles for mixed waters. Average dimensionless C profiles for each trophic category (a), rescaled C profiles for categories M1 to M3 (b), and rescaled C profiles for categories M4 and M5 (c). For rescaled profiles of each trophic category, the position of Z_{eu} is identified by an arrow associated with the symbol relevant to the trophic category.

Figure 6: Comparison of parameterizations developed in the present study and in MB89 to retrieve the C vertical profile from C_{surf}. Three criteria are compared. Integrated content within the euphotic layer $\langle C \rangle_{Zeu}$ (a), maximal concentration C_{max} (b) and depth of the maximal concentration Z_{max} (c).

Figure 7: Change in the size structure of the phytoplankton community as a function of C_{surf} for stratified waters. Average proportion (%) of micro-C, nano-C and pico-C, within the euphotic layer (a). Average integrated content of micro-C, nano-C, pico-C and C within the euphotic layer (b). Average proportion (%) of micro-C, nano-C and pico-C, within the surface layer (c). The vertical bars represent ± 1 sd around mean values.

Figure 8: Average (a-c) and modeled (d-f) dimensionless vertical profiles of micro-C, nano-C and pico-C in stratified waters.

Figure 9: Average rescaled vertical profiles of micro-C (a-b), nano-C (c-d) and pico-C (e-f) in stratified waters. For each trophic category, the position of Z_{eu} is identified by an arrow associated with the symbol relevant to the trophic category.

Figure 10: Change in the size structure of the phytoplankton community as a function of C_{surf} for mixed waters. Average proportion (%) of micro-C, nano-C and pico-C, within the euphotic layer for global mixed waters, southern waters excluded (a) and for southern mixed waters (b). The vertical bars represent ± 1 sd around mean values.

Figure 11: Average dimensionless vertical profiles of micro-C, nano-C, and pico-C in mixed waters. Data are split into global mixed waters (excluding southern waters) (a-c), and southern mixed waters (d-f).

Figure 12: Comparison of measured and computed characteristics of vertical profiles of micro-C, nano-C and pico-C. Integrated content within the euphotic zone for stratified and mixed waters (a-c); maximal concentration for stratified waters (d-f); depth of the maximal concentration for stratified waters (g-i).

Figure 13: Comparison of measured and computed micro-C (a), nano-C (b) and pico-C (c) for surface waters.

Figure 14: Global distribution of the size structure of the phytoplankton community for June 2000 (SeaWiFS composite). Proportion (%) of micro-C, nano-C and pico-C (a-c), and integrated content within the euphotic layer (mg m^{-2}) of micro-C, nano-C and pico-C (d-f).

Table 1: Symbols used in the present study and their significance.

SYMBOL	SIGNIFICANCE
C	Total chlorophyll *a* (divinyl-chlorophyll *a* + monovinyl-chlorophyll *a* + chlorophyllid *a*) concentration, mg m^{-3}
PAR	Photosynthetically available radiation, W m^{-2}
K	Attenuation coefficient for downwelling PAR irradiance, m^{-1}
Z_{pd}	Depth of the surface layer, or penetration depth (= 1 / K), m
C_{surf}	Average integrated C within the layer 0-Z_{pd}, mg m^{-3}
z	Geometrical depth, m
Z_{eu}	Depth of the euphotic layer, defined as the depth where the PAR is reduced to 1% of its surface value, m
ζ	Depth normalized with respect to Z_{eu}: $\zeta = z / Z_{eu}$, dimensionless
$\langle P \rangle_{Zeu}$	Column-integrated content of pigment P within the euphotic layer, mg m^{-2}
$\overline{C}_{Zeu}$	Average column-integrated content of C within the euphotic layer, mg m^{-3}
$P(\zeta)$	Concentration of pigment P, at (the dimensionless depth) ζ, mg m^{-3}
$p(\zeta)$	Concentration of pigment P normalized with respect to $\overline{C}_{Zeu}$, $p(\zeta) = P(\zeta) / \overline{C}_{Zeu}$, dimensionless
micro-C	Total chlorophyll *a* concentration associated to micro-phytoplankton, mg m^{-3}
nano-C	Total chlorophyll *a* concentration associated to nano-phytoplankton, mg m^{-3}
pico-C	Total chlorophyll *a* concentration associated to pico-phytoplankton, mg m^{-3}
[%micro]	Proportion of biomass associated to micro-phytoplankton (%)
[%nano]	Proportion of biomass associated to nano-phytoplankton (%)
[%pico]	Proportion of biomass associated to pico-phytoplankton (%)
Z_m	Depth of the mixed layer, m
P_{max}	Maximum concentration on a considered P vertical profile, mg m^{-3}
Z_{max}	Depth of P_{max}, m

Table 2: Biomarker pigments used in the present study and their taxonomic significance, associated size classe, and corresponding C to P ratio.

Diagnostic pigments	Abbreviations	Taxonomic significance	Size range	Phytoplankton size class	$\langle C \rangle_{Zeu} : \langle P \rangle_{Zeu}$
Fucoxanthin	Fuco	diatoms	> 20 μm	micro-phytoplankton	1.41
Peridinin	Peri	dinoflagellates			1.41
19′-hexanoyloxyfucoxanthin	19′-HF	chromophytes nanoflagellates	2-20 μm	nano-phytoplankton	1.27
19′-butanoyloxyfucoxanthin	19′-BF	chromophytes nanoflagellates			0.35
Alloxanthin	Allo	cryptophytes			0.60
chlorophyll *b* + divinyl-chlorophyll *b*	TChl*b*	green flagellates and prochlorophytes	< 2 μm	pico-phytoplankton	1.01
Zeaxanthin	Zea	cyanobacteria and prochlorophytes			0.86

Table 3: Statistical relationships between the average total chlorophyll *a* concentration within the surface layer, C_{surf} (mg m^{-3}), and the total chlorophyll *a* integrated content within the euphotic layer, $\langle C \rangle_{Zeu}$ (mg m^{-2}). The relationships obtained by MB89 are reproduced in italic.

Stratified waters (n = 1821 profiles)			
For $C_{surf} \leq 1$ mg m^{-3}			
$\langle C \rangle_{Zeu} = 36.1\, C_{surf}^{0.357}$	(8a)	($r^2 = 0.73$, $p < 0.001$)	$\langle C \rangle_{Zeu} = 38.0\, C_{surf}^{0.425}$
For $C_{surf} > 1$ mg m^{-3}			
$\langle C \rangle_{Zeu} = 37.7\, C_{surf}^{0.615}$	(8b)	($r^2 = 0.70$, $p < 0.001$)	$\langle C \rangle_{Zeu} = 40.2\, C_{surf}^{0.507}$
Mixed waters (n = 598 profiles)			
$\langle C \rangle_{Zeu} = 42.1\, C_{surf}^{0.538}$	(8c)	($r^2 = 0.95$, $p < 0.001$)	$\langle C \rangle_{Zeu} = 38.0\, C_{surf}^{0.551}$

Table 4: Trophic categories defined with respect to the average total chlorophyll *a* concentration within the surface layer, C_{surf}, and the associated relevant parameters. The parameters are derived from the calculations performed on the complete database #1, and are presented as averages and standard deviations (the latter are shown in parentheses).

	Stratified waters									Mixed waters				
Trophic classes	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**M1**	**M2**	**M3**	**M4**	**M5**
C_{surf} range, (mg m^{-3})	<0.04*	0.04-0.08	0.08-0.12	0.12-0.2	0.2-0.3	0.3-0.4	0.4-0.8	0.8-2.2	2.2-4**	<0.4§	0.4-0.8	0.8-1	1-4	>4¶
Number of profiles	109	268	269	320	287	180	260	110	18	155	153	53	182	55
Average C_{surf} (mg m^{-3})	0.032 (0.005)	0.062 (0.012)	0.098 (0.012)	0.158 (0.023)	0.244 (0.030)	0.347 (0.028)	0.540 (0.106)	1.235 (0.403)	2.953 (0.520)	0.244 (0.092)	0.592 (0.112)	0.885 (0.051)	1.881 (0.753)	6.320 (2.916)
Average $\overline{C}_{Zeu}$ (mg m^{-3})	0.0910 (0.025)	0.151 (0.067)	0.185 (0.088)	0.250 (0.144)	0.338 (0.152)	0.410 (0.153)	0.578 (0.229)	1.206 (0.526)	2.950 (1.191)	0.280 (0.130)	0.591 (0.175)	0.872 (0.189)	2.059 (0.996)	7.574 (3.700)
Average $\langle C \rangle_{Zeu}$ (mg m^{-3})	10.54 (1.84)	14.15 (3.31)	15.98 (3.29)	18.79 (4.08)	22.09 (4.99)	24.70 (4.64)	29.72 (5.88)	44.05 (10.46)	71.98 (15.28)	19.90 (4.70)	30.27 (4.73)	37.57 (4.44)	58.64 (15.30)	120.00 (26.75)
Average Z_{eu} (m)	119.1 (12.2)	99.9 (15.4)	91.0 (11.8)	80.2 (12.6)	70.3 (11.9)	63.4 (9.3)	54.4 (8.2)	39.8 (8.0)	26.1 (4.5)	77.1 (14.3)	53.2 (6.8)	44.0 (4.6)	31.5 (6.8)	16.9 (2.4)

* Minimum value 0.015 mg m^{-3} ** Maximum value 3.97 mg m^{-3}
§ Minimum value 0.047 mg m^{-3} ¶ Maximum value 23.9 mg m^{-3}

Table 5: Values of the 5 parameters of equation (7), obtained for the average dimensionless vertical profiles of C, micro-C, nano-C, and pico-C, for each trophic class of the stratified waters (S1 to S9).

	C					micro-C					nano-C					pico-C				
Trophic class	C_b	s	C_{max}	ζ_{max}	$\Delta\zeta$	C_b	s	C_{max}	ζ_{max}	$\Delta\zeta$	C_b	s	C_{max}	ζ_{max}	$\Delta\zeta$	C_b	s	C_{max}	ζ_{max}	$\Delta\zeta$
S1	0.471	0.135	1.572	0.969	0.393	0.036	0.020	0.122	1.012	0.532	0.138	0.033	0.764	0.980	0.451	0.222	0.114	0.906	0.970	0.352
S2	0.533	0.172	1.194	0.921	0.435	0.071	0.020	0.173	0.885	0.406	0.129	0.014	0.589	0.899	0.454	0.242	0.109	0.627	0.977	0.427
S3	0.428	0.138	1.015	0.905	0.630	0.076	0.021	0.126	0.835	0.424	0.142	0	0.463	0.872	0.526	0.254	0.099	0.437	0.969	0.634
S4	0.570	0.173	0.766	0.814	0.586	0.071	0.021	0.160	0.776	0.546	0.192	0.037	0.400	0.782	0.535	0.271	0.100	0.255	0.858	0.637
S5	0.611	0.214	0.676	0.663	0.539	0.145	0.050	0.163	0.700	0.479	0.188	0.055	0.418	0.650	0.640	0.159	0.052	0.176	0.574	0.650
S6	0.390	0.109	0.788	0.521	0.681	0.173	0.044	0.161	0.600	0.508	0.331	0.132	0.294	0.501	0.516	0.176	0.071	0.129	0.458	0.626
S7	0.569	0.183	0.608	0.452	0.744	0.237	0.077	0.158	0.521	0.543	0.201	0.084	0.350	0.402	0.724	0.009	0	0.251	0.239	0.943
S8	0.835	0.298	0.382	0.512	0.625	0.331	0.105	0.278	0.451	0.746	0.227	0.081	0.198	0.181	0.690	0.094	0.040	0.109	0.187	0.618
S9	0.188	0	0.885	0.378	1.081	0.891	0.302	0.000	0.277	1.014	0.171	0	0.088	0.375	0.352	0.051	0.023	0.000	0.052	0.417

Table 6: Average proportions (%) of micro-C, nano-C, and pico-C within the euphotic layer, for the average C_{surf} values of each trophic category for mixed waters (M1 to M5).

	Trohic class	Average C_{surf} (mg m^{-3})	%micro	%nano	%pico
Global mixed waters (southern waters excluded)	M1	0.234	17.7	51.0	31.3
	M2	0.593	24.0	50.1	25.9
	M3	0.891	27.5	58.0	14.4
	M4	1.540	51.9	38.7	9.4
	M5	7.964	90.6	5.6	3.8
Southern mixed waters	M1	0.345	54.7	34.9	10.4
	M2	0.605	50.6	44.4	5.0
	M3	0.889	44.3	50.6	5.0
	M4	1.956	41.0	56.6	5.0
	M5	5.755	13.3	85.7	1.0

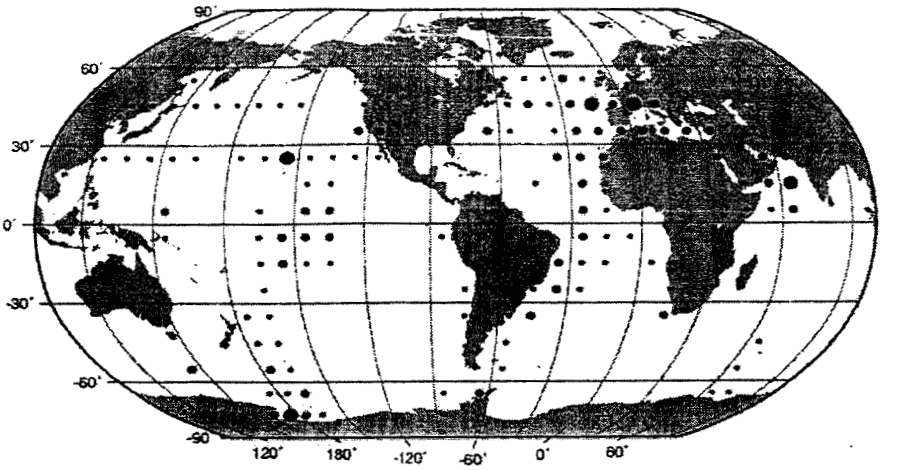


Figure 1: Geographic location of the 2419 stations used in the present study. The size of the black dots indicates the number of stations within a square of 10° longitude by 10° latitude. The small dot corresponds to less than 10 stations; the medium to 10-100 stations; and the large to more than 100 stations.

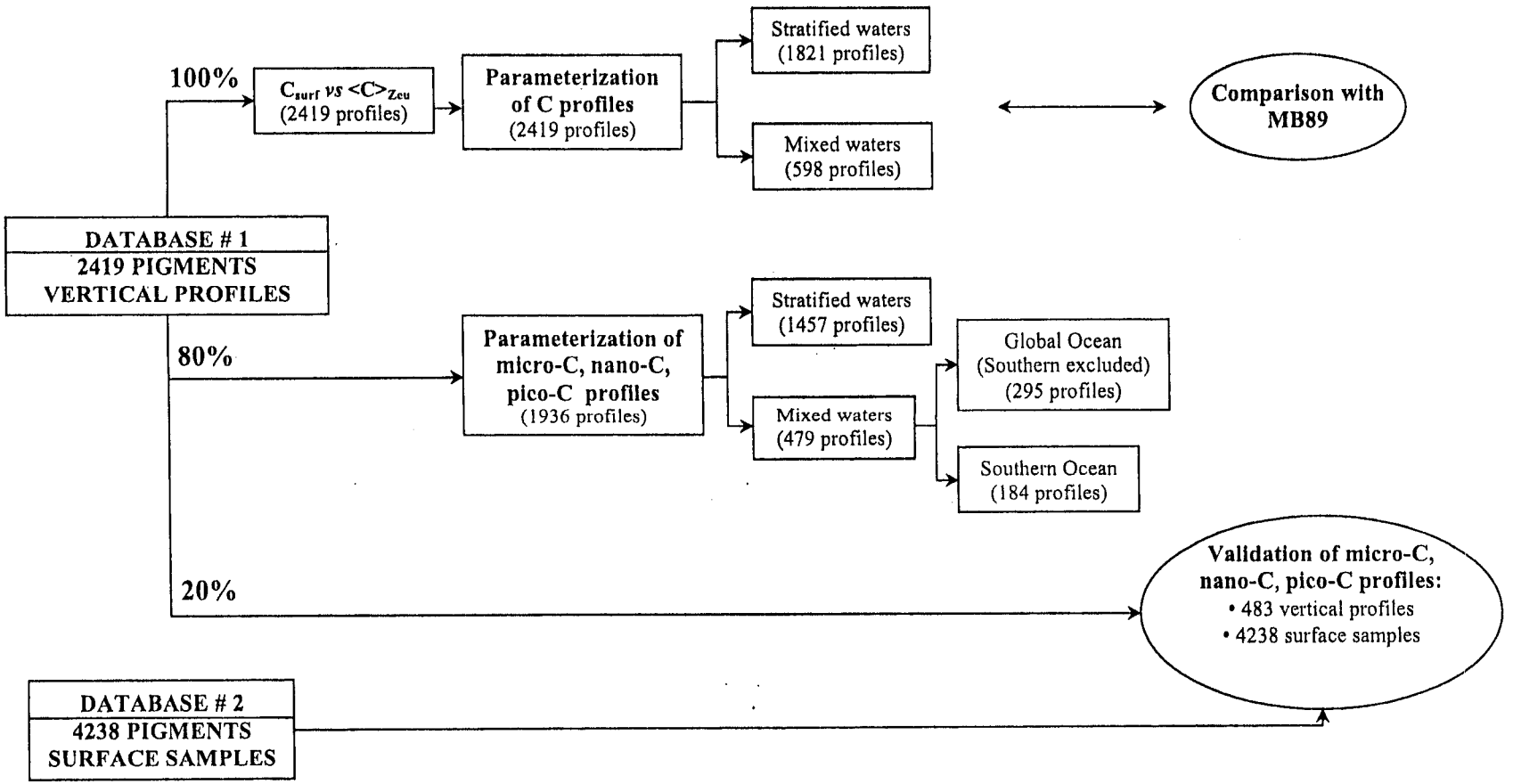


Figure 2: Flowchart detailing the partition and the use of the two data sets for data analysis as well as for validation of the developed parameterization.

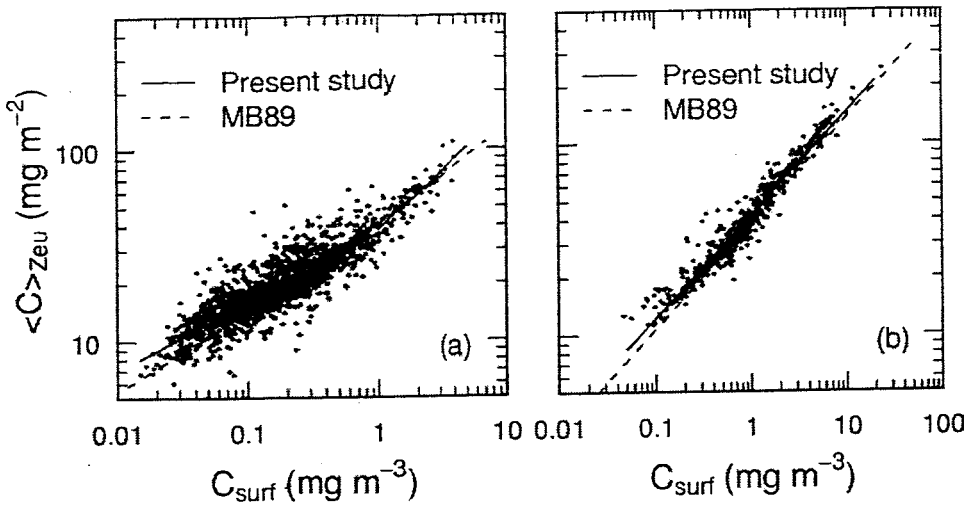


Figure 3: Relationship between the average total chlorophyll *a* concentration within the surface layer, C_{surf}, and the total chlorophyll *a* integrated content within the euphotic layer, $\langle C \rangle_{Zeu}$, for stratified (a) and mixed (b) waters. The solid lines represent equations (8) while the dashed lines correspond to the regression lines models of MB89.

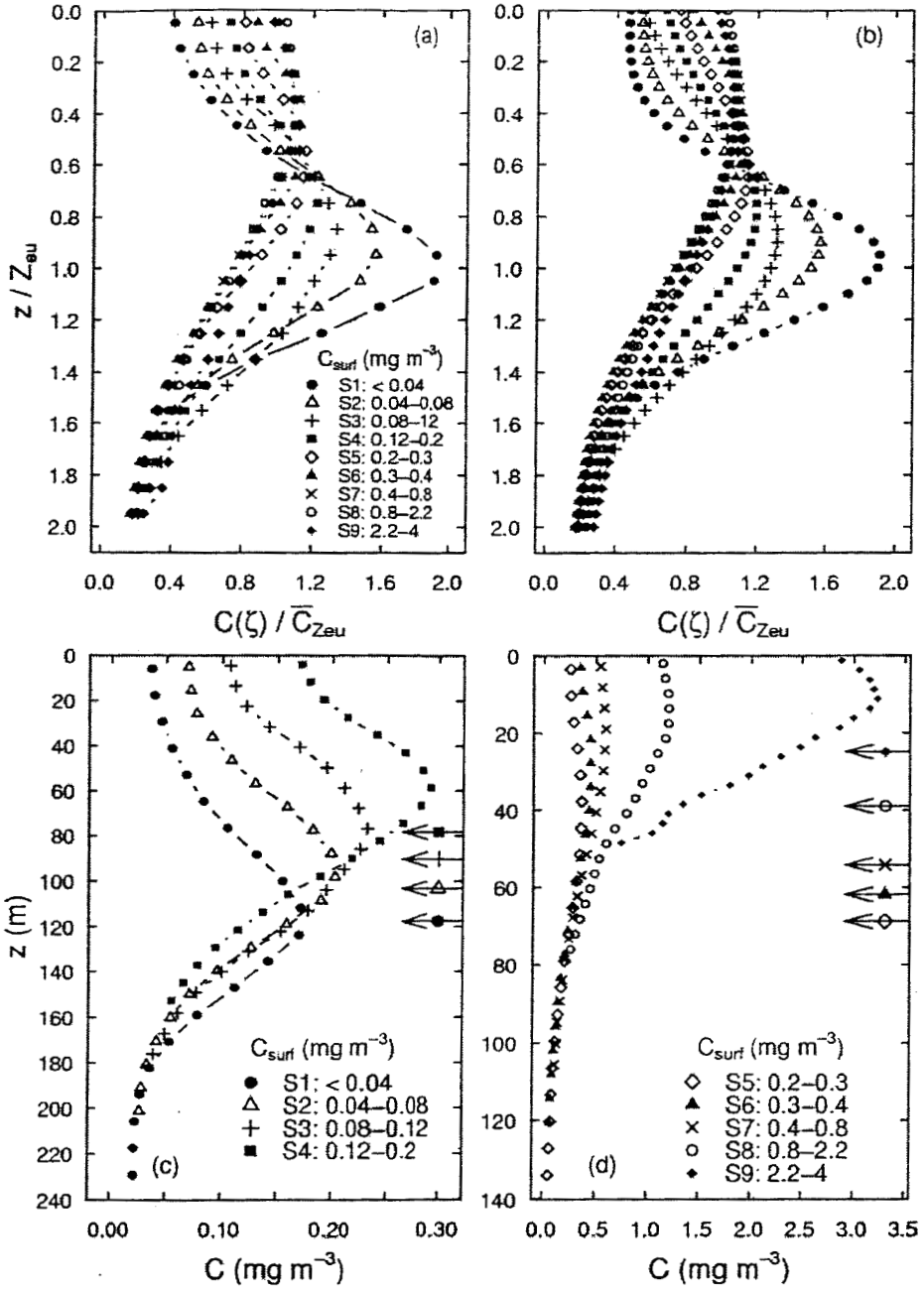


Figure 4: C vertical profiles for stratified waters. Average dimensionless C profiles for each trophic category (a), modeled dimensionless C profiles for each trophic category (b), rescaled C profiles for categories S1 to S4 (c), and rescaled C profiles for categories S5 to S9 (d). For rescaled profiles of each trophic category, the position of Z_{eu} is identified by an arrow associated with the symbol relevant to the trophic category.

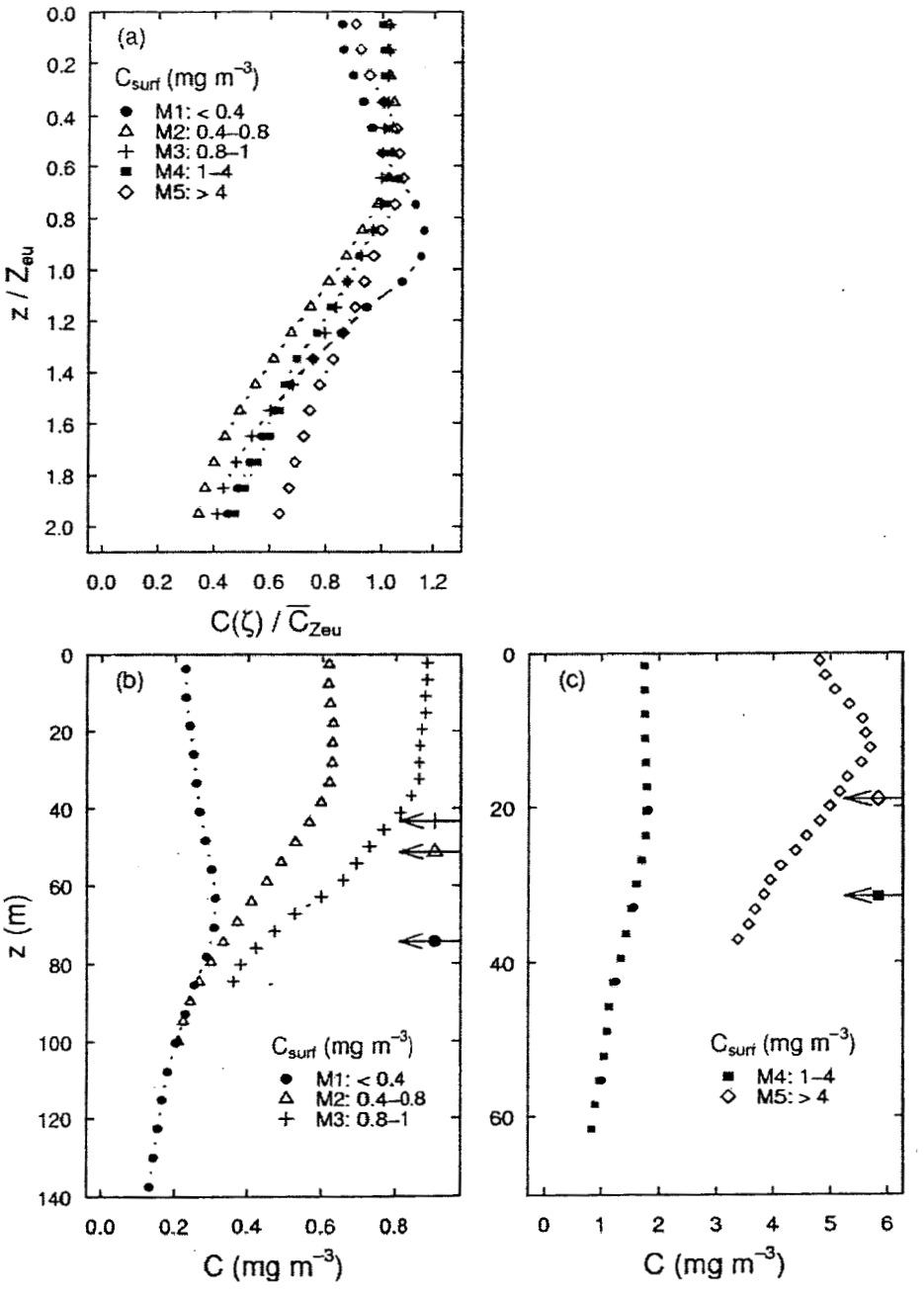


Figure 5: C vertical profiles for mixed waters. Average dimensionless C profiles for each trophic category (a), rescaled C profiles for categories M1 to M3 (b), and rescaled C profiles for categories M4 and M5 (c). For rescaled profiles of each trophic category, the position of Z_{eu} is identified by an arrow associated with the symbol relevant to the trophic category.

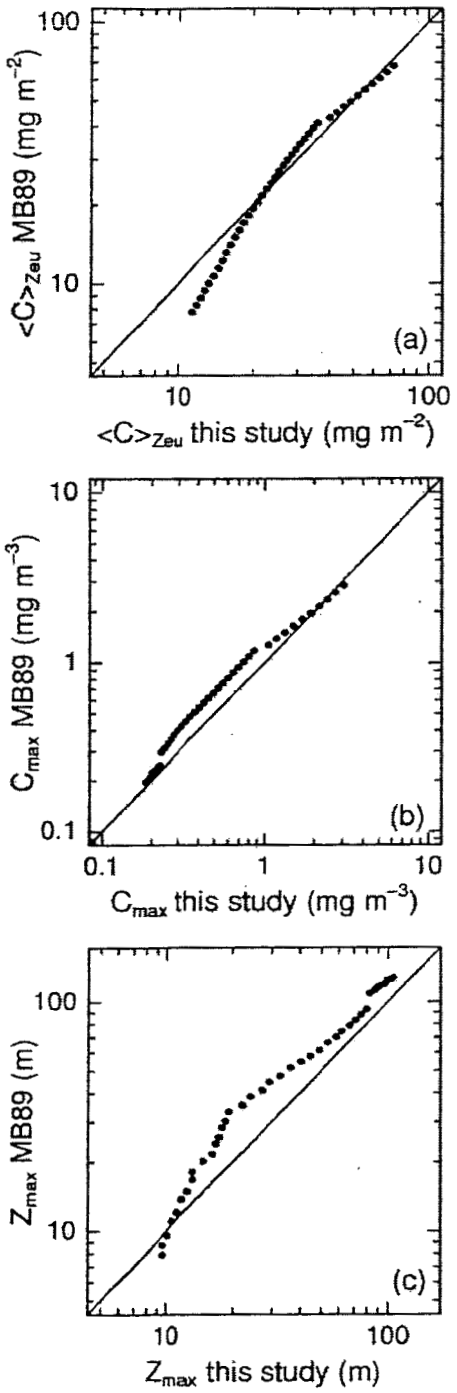


Figure 6: Comparison of parameterizations developed in the present study and in MB89 to retrieve the C vertical profile from C_{surf}. Three criteria are compared. Integrated content within the euphotic layer $\langle C \rangle_{Zeu}$ (a), maximal concentration C_{max} (b) and depth of the maximal concentration Z_{max} (c).

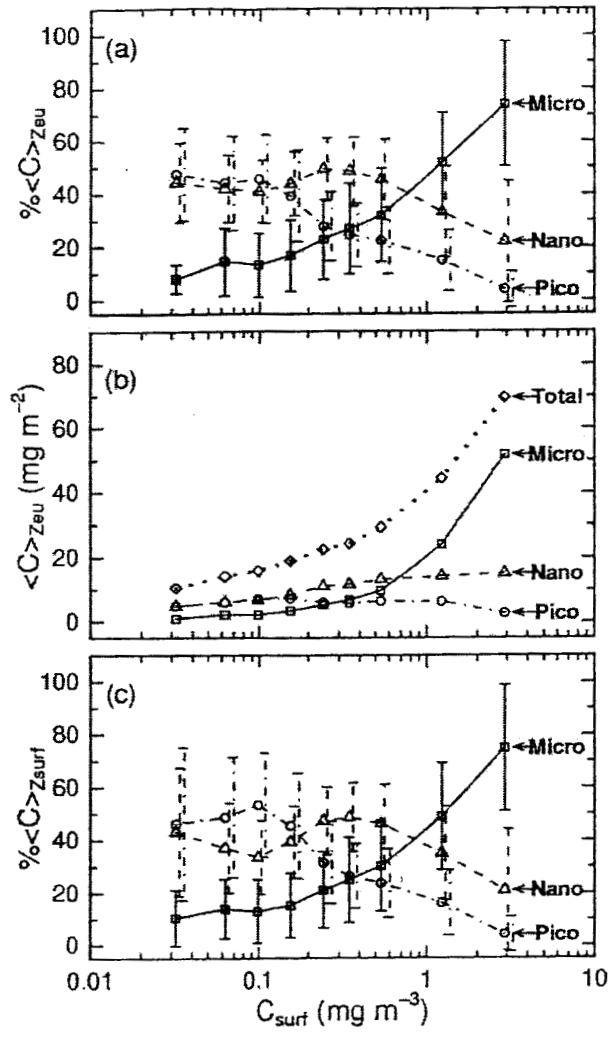


Figure 7: Change in the size structure of the phytoplankton community as a function of C_{surf} for stratified waters. Average proportion (%) of micro-C, nano-C and pico-C, within the euphotic layer (a). Average integrated content of micro-C, nano-C, pico-C and C within the euphotic layer (b). Average proportion (%) of micro-C, nano-C and pico-C, within the surface layer (c). The vertical bars represent ± 1 sd around mean values.

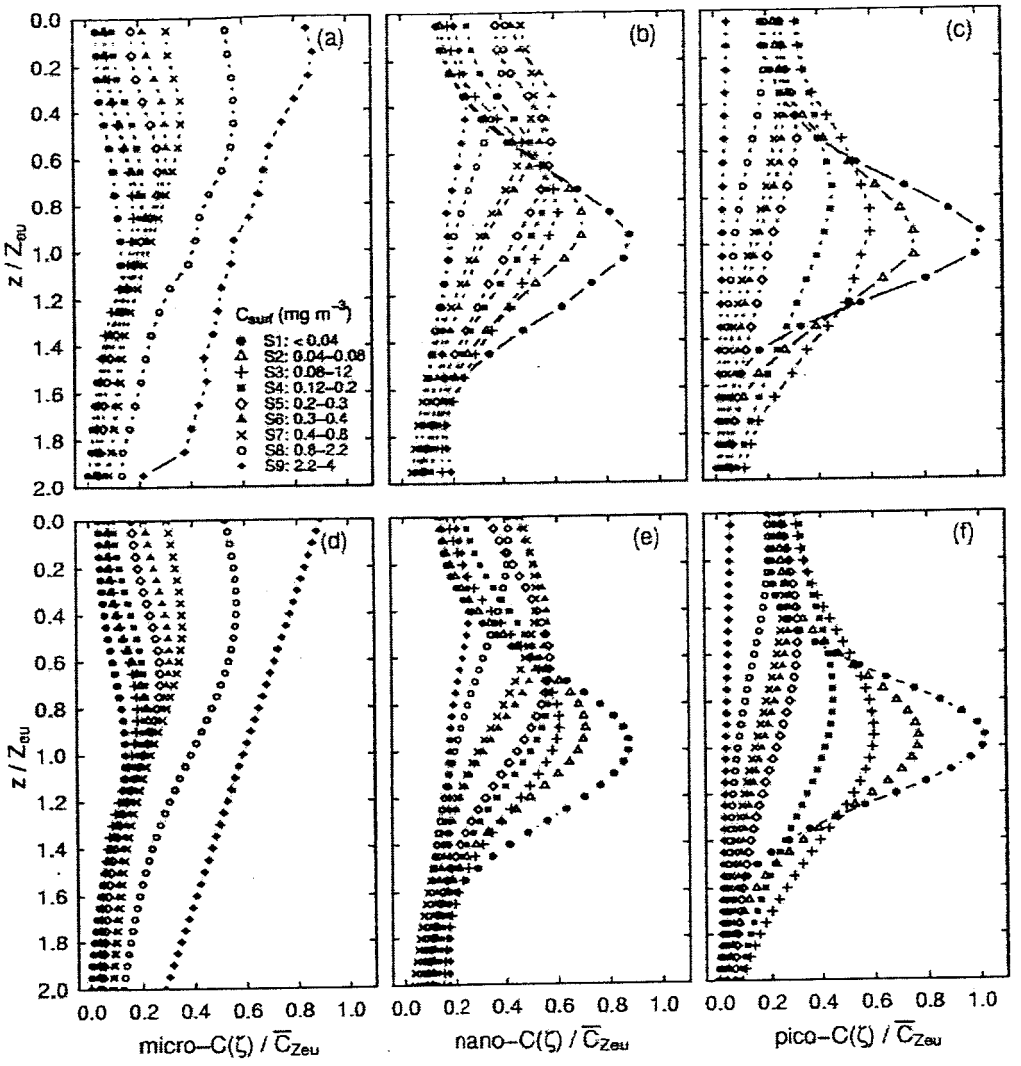


Figure 8: Average (a-c) and modeled (d-f) dimensionless vertical profiles of micro-C, nano-C and pico-C in stratified waters.

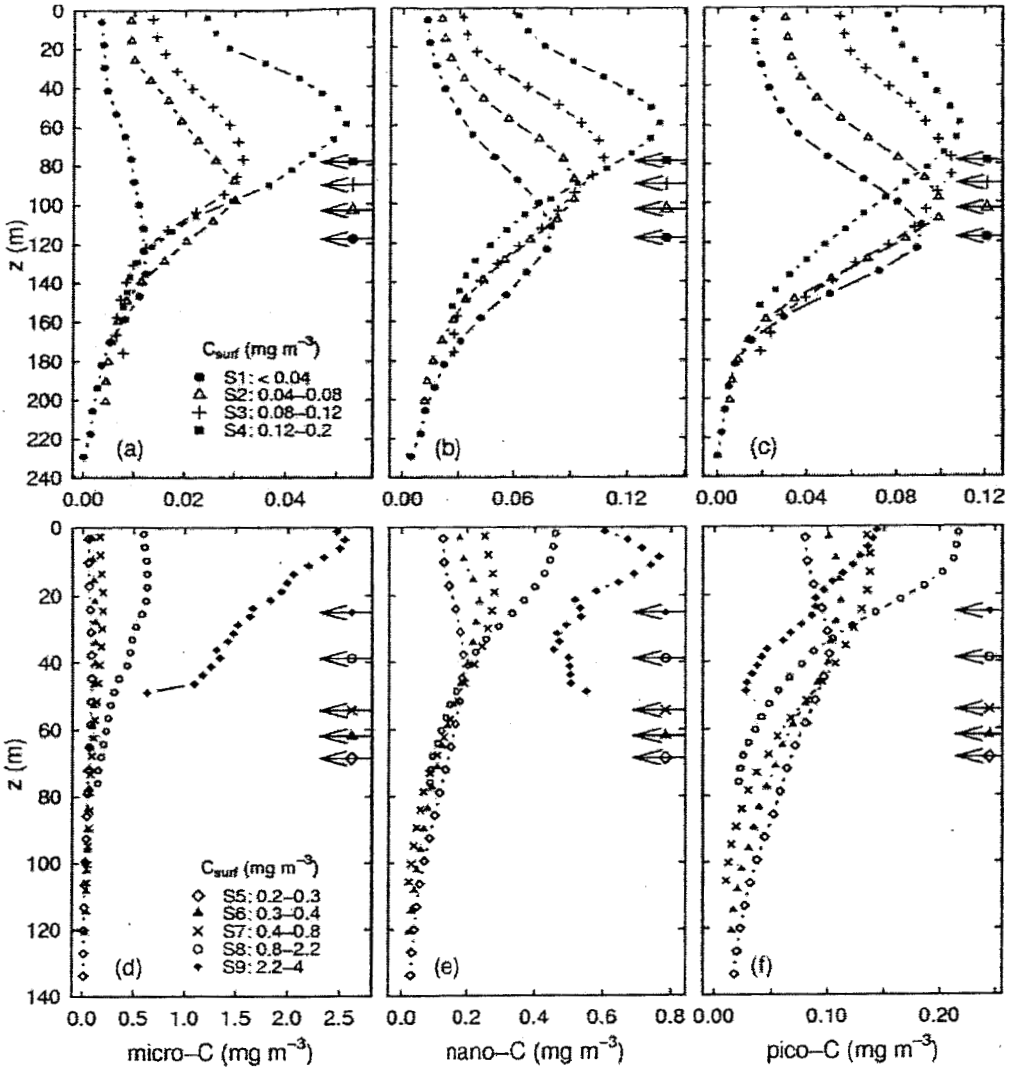


Figure 9: Average rescaled vertical profiles of micro-C (a-b), nano-C (c-d) and pico-C (e-f) in stratified waters. For each trophic category, the position of Z_{eu} is identified by an arrow associated with the symbol relevant to the trophic category.

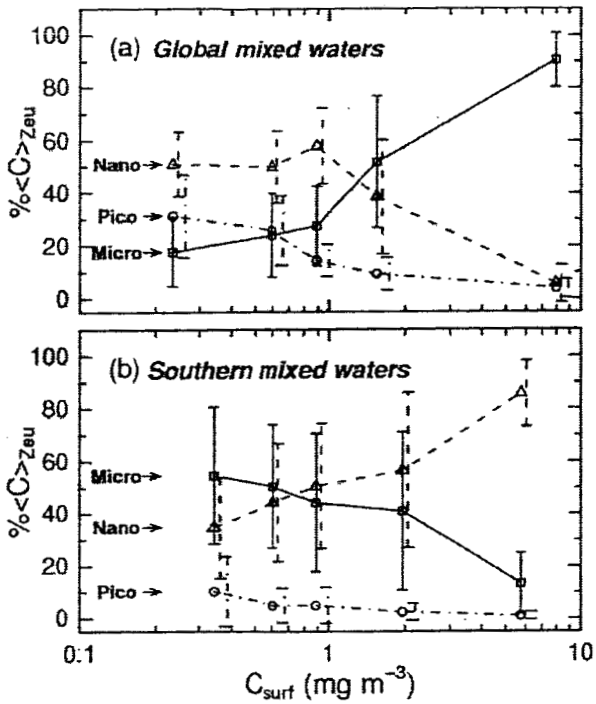


Figure 10: Change in the size structure of the phytoplankton community as a function of C_{surf} for mixed waters. Average proportion (%) of micro-C, nano-C and pico-C, within the euphotic layer for global mixed waters, southern waters excluded (a) and for southern mixed waters (b). The vertical bars represent ± 1 sd around mean values.

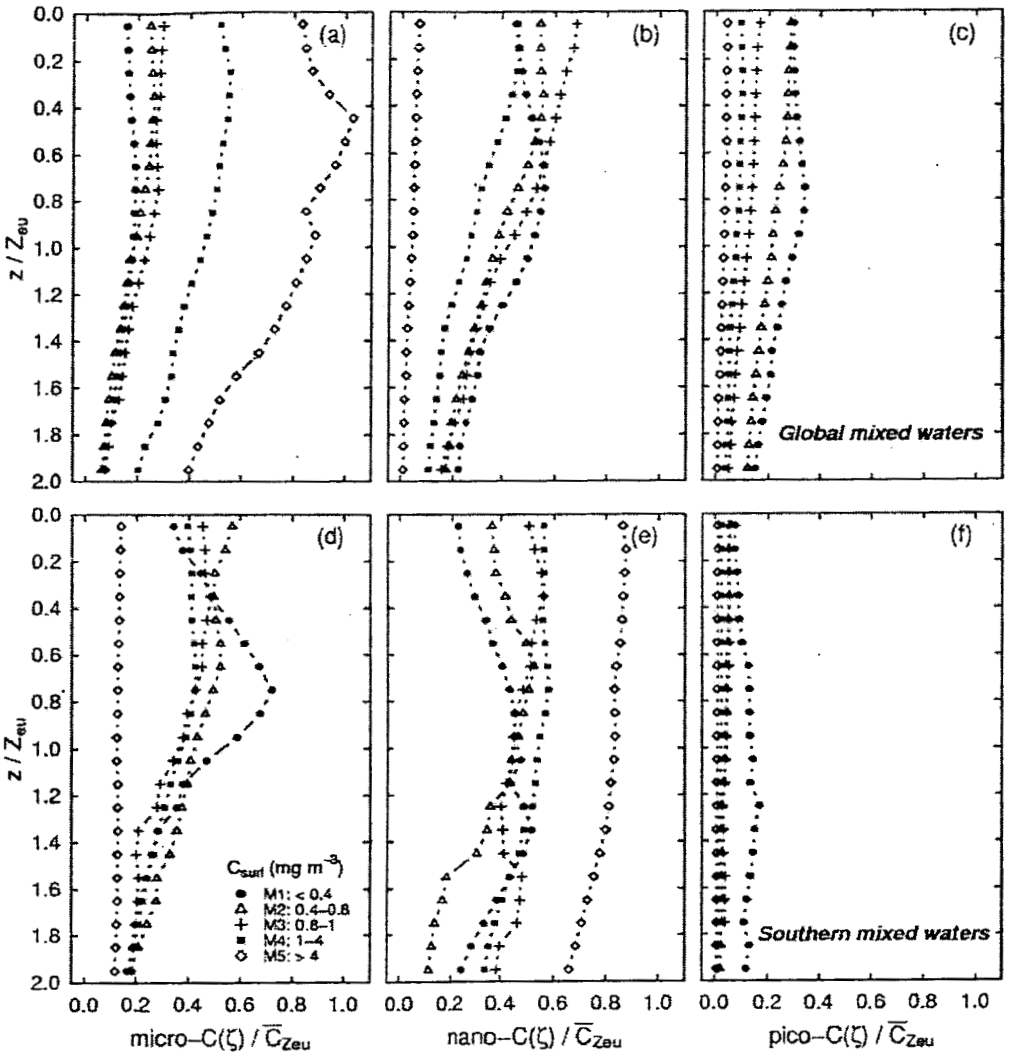


Figure 11: Average dimensionless vertical profiles of micro-C, nano-C, and pico-C in mixed waters. Data are split into global mixed waters (excluding southern waters) (a-c), and southern mixed waters (d-f).

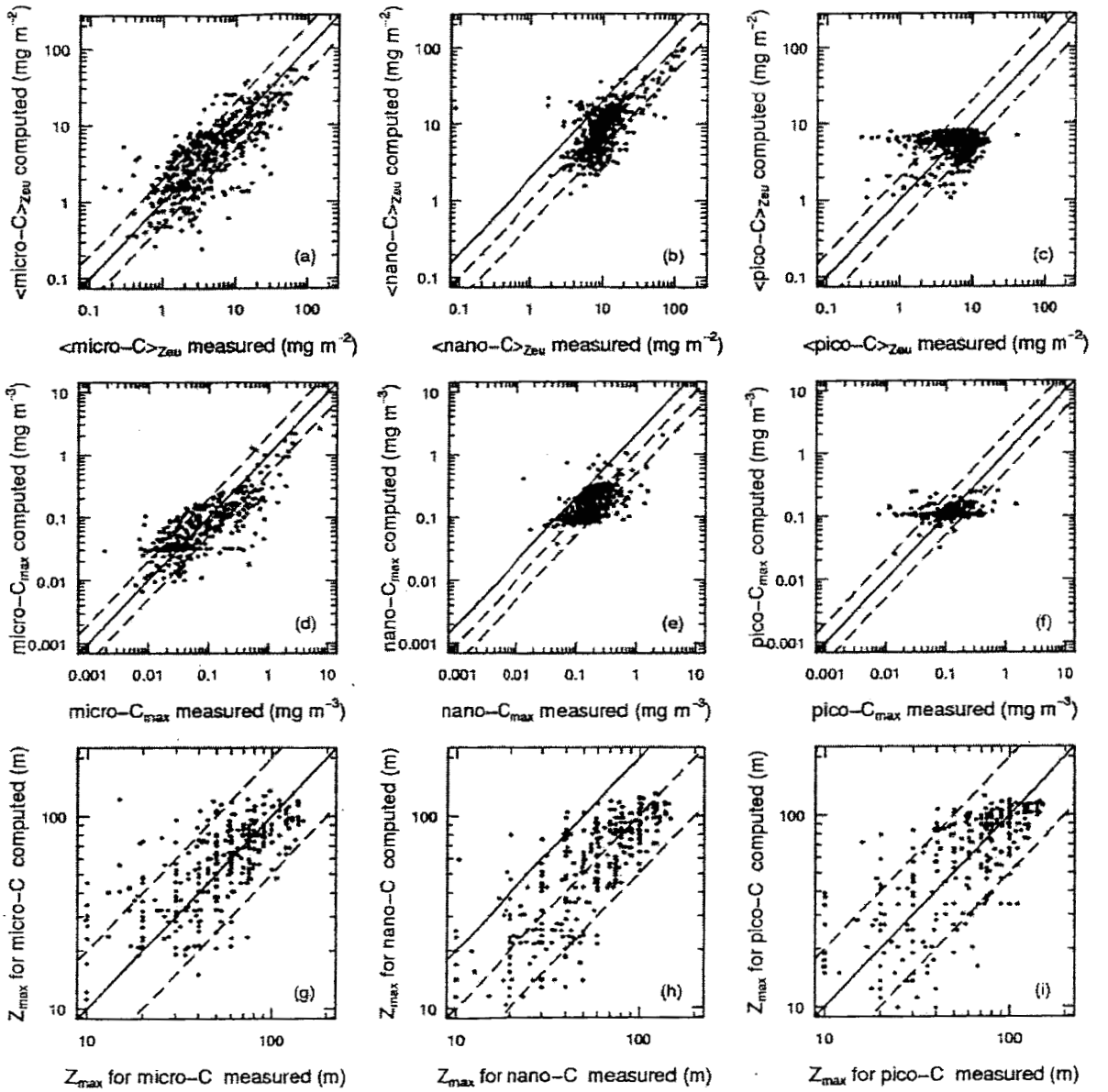


Figure 12: Comparison of measured and computed characteristics of vertical profiles of micro-C, nano-C and pico-C. Integrated content within the euphotic zone for stratified and mixed waters (a-c); maximal concentration for stratified waters (d-f); depth of the maximal concentration for stratified waters (g-i).

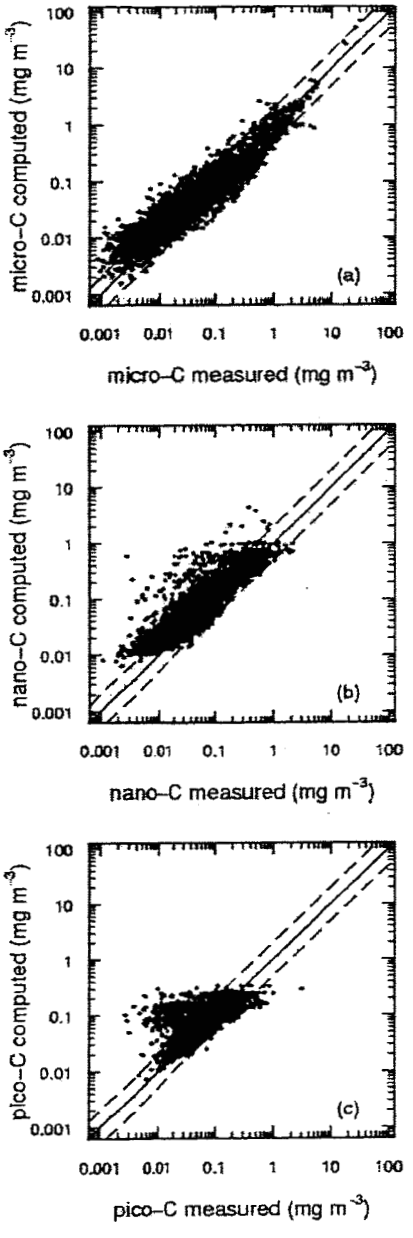


Figure 13: Comparison of measured and computed micro-C (a), nano-C (b) and pico-C (c) for surface waters.

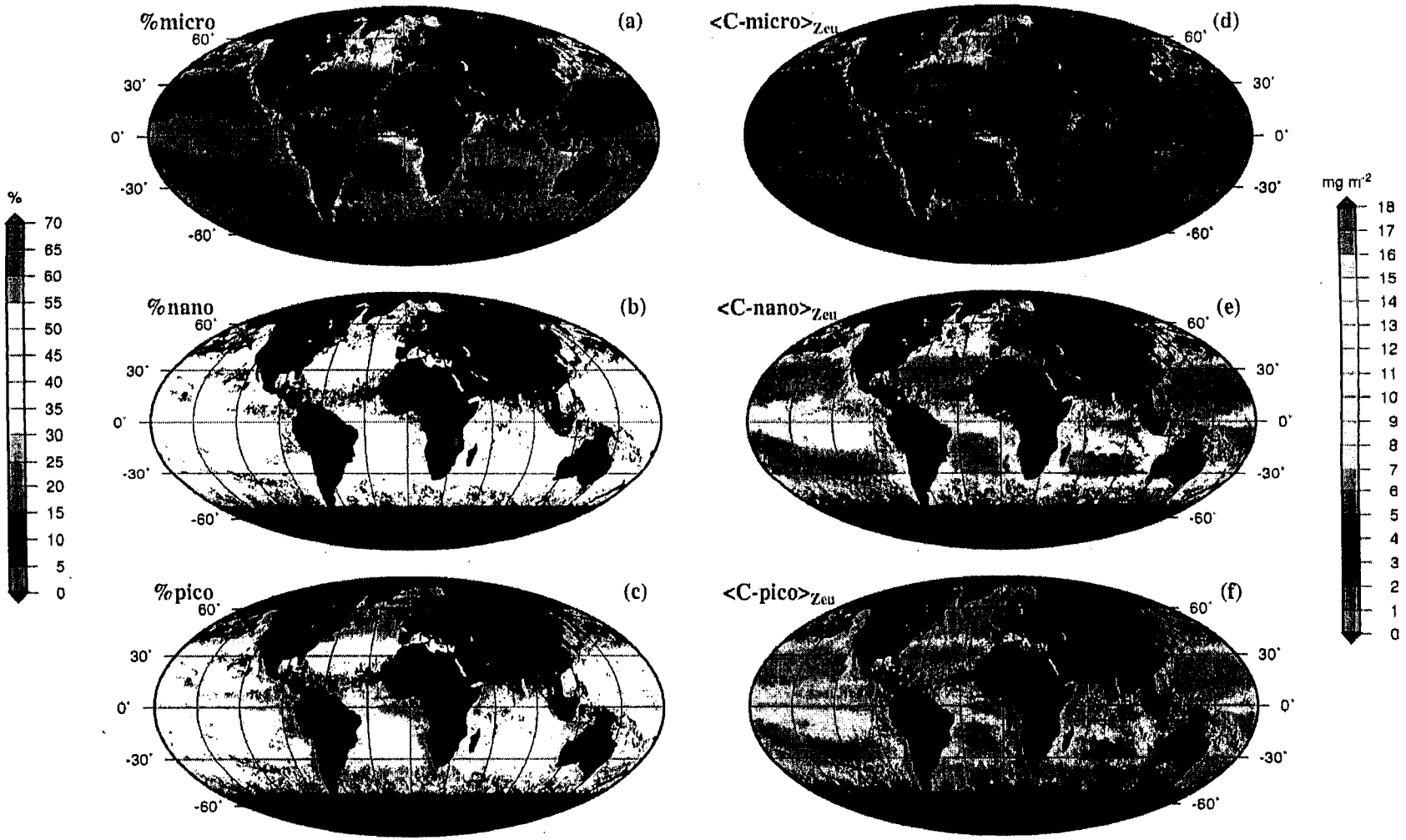


Figure 14: Global distribution of the size structure of the phytoplankton community for June 2000 (SeaWiFS composite). Proportion (%) of micro-C, nano-C and pico-C (a-c), and integrated content within the euphotic layer (mg m^{-2}) of micro-C, nano-C and pico-C (d-f).

Electronic supplement

Appendix A: Summary of the data collection activities for both databases used in the present study.

A1: Origin of the data included in database #1

Project and Cruise	Year	Location	Source	Ref	n
ACSOE DI234	1998	NE Atlantic	http://www.bodc.ac.uk		8
AESOPS Kiwi 6	1997	Ross Sea	http://usjgofs.whoi.edu/jg/dir/jgofs/		4
AESOPS Kiwi 7	1997-1998	Ross Sea	-		11
AESOPS Kiwi 8	1998	Ross Sea	-		5
AESOPS Kiwi 9	1998	Ross Sea	-		8
AESOPS NBP 96-4A	1996	Ross Sea	-		1
AESOPS NBP 97-1	1997	Ross Sea	-		17
AESOPS NBP 97-8	1997	Ross Sea	-		24
ALMOFRONT 1	1991	Alboran Sea, Mediterranean Sea	http://www.obs-vlfr.fr/cd_rom_dmtt/start_france_jgofs.htm	1	22
ALMOFRONT 2	1998	Alboran Sea, Mediterranean Sea	-	-	39
AMT-1	1995	Trans Atlantic	S. Hooker, J. Aiken, and G. Westbrook		14
AMT-2	1996	Trans Atlantic	G. Moore, S. Hooker, G. Westbrook	2	14
AMT-3	1996	Trans Atlantic	S. Hooker, J. Aiken, and R. Barlow	3	25
AMT-4	1997	Trans Atlantic	S. Hooker, J. Aiken, and L. Van Heukelem	2	27
AMT-5	1997	Trans Atlantic	S. Hooker, J. Aiken, and R. Barlow	-	26
AMT-6	1998	Trans Atlantic	-	3	22
AMT-7	1998	Trans Atlantic	S. Hooker, J. Aiken, and L. Dransfield		25
AMT-8	1999	Trans Atlantic	-		23
AMT-11	2000	Trans Atlantic	J. Aiken, C. Omachi, and M. Zapata		4
ANTARES 2	1994	Indian Sector of the Southern Ocean	http://www.obs-vlfr.fr/cd_rom_dmtt/start_france_jgofs.htm	4	14
ARABESQUE DI210	1994	Arabian Sea	C. Trees and R. Barlow / http://www.bodc.ac.uk	5	29
ARABESQUE DI212	1994	Arabian Sea	R. Barlow / http://www.bodc.ac.uk	-	26
Arabian Sea TTN-043	1995	Arabian Sea	http://usjgofs.whoi.edu/jg/dir/jgofs/arabian/		63
Arabian Sea TTN-045	1995	Arabian Sea	-		45
Arabian Sea TTN-049	1995	Arabian Sea	-		63
Arabian Sea TTN-050	1995	Arabian Sea	-		44
Arabian Sea TTN-053	1995	Arabian Sea	-		45
Arabian Sea TTN-054	1995	Arabian Sea	-		56
BAF 89/3	1989	N Atlantic	http://www.pangaea.de/PangaVista?header=ProjectPageHeader&headerparam=JGOFS@Data&query=@Ref23423		5
BATS Time-series	1989-1998	N Atlantic	http://www.bbsr.edu/cintoo/bats/bats.html		94

BENCAL	2002	Benguela Upwelling	J. Ras		11
BOFS DI183	1989	NE Atlantic	http://www.bodc.ac.uk		2
BOFS CD46	1990	NE Atlantic	R. Barlow / http://www.bodc.ac.uk	6	1
BOFS DI191	1990	NE Atlantic	-		2
BOFS STERNA JCR2	1992	Bellingshausen Sea	R. Barlow / http://www.bodc.ac.uk	7	5
CalCOFI 9308	1993	Off California Coast	http://seabass.gsfc.nasa.gov/		31
CalCOFI 9310	1993	Off California Coast	-		20
CalCOFI 9410	1994	Off California Coast	-		20
CalCOFI 9507	1995	Off California Coast	-		19
CalCOFI 9602	1996	Off California Coast	-		2
CATCH	1997	N Atlantic	K. Oubelkheir		27
COQFEB99	1999	Off Chile coast	V. Stuart and E. Head		1
DCM	1996	N Atlantic	G. Kraay and M. Veldhuis	8	3
DYFAMED Time-series	1991-1997	Mediterranean Sea	http://www.obs-vlfr.fr/cd_rom_dmtt/start_france_jgofs.htm		66
DYNAPROC	1993	Mediterranean Sea	http://www.obs-vlfr.fr/cd_rom_dmtt/start_france_jgofs.htm	9	63
EqPac TT-007	1992	Equatorial Pacific	http://usjgofs.whoi.edu/jg/dir/jgofs/eqpac	10	39
EqPac TT-008	1992	Equatorial Pacific	-	-	32
EqPac TT-011	1992	Equatorial Pacific	-	-	39
EqPac TT-012	1992	Equatorial Pacific	-	-	33
EUMELI 3	1991	NE Subtropical Atlantic	http://www.obs-vlfr.fr/cd_rom_dmtt/start_france_jgofs.htm	11	16
EUMELI 4	1992	NE Subtropical Atlantic	-	-	25
FR 08/90	1990	W Equatorial Pacific	http://www.marine.csiro.au/datacentre/JGOFSweb/cmr_jgofs.htm	12	4
FR 05/92	1992	W Equatorial Pacific	-	13, 14	16
HOT Time-series	1988-1999	N Pacific	http://hahana.soest.hawaii.edu/hot/hot_jgofs.html		93
ICECOLORS 90	1990	Bellingshausen Sea	R. Bidigare	15	54
IRONEX I 93	1993	S Pacific	-	15	6
JG96MAY	1996	Labrador Sea	V. Stuart and E. Head	16	3
JG97MAY	1997	Labrador Sea	-		1
JGOFS-Canada	1996	N Pacific	S. Roy and D. Thibault		9
LABSEA2000	2000	Labrador Sea	-		3
MARE I	2000	South Atlantic	G. Kraay and M. Veldhuis		1
METEOR	1995	Mediterranean Sea	F. Vidussi	17	70
MINOS	1996	Mediterranean Sea	-		69
MIRC98OCT	1998	Off Chile coast	V. Stuart and E. Head	18	8
MLML 91	1991	Off Iceland coast	R. Bidigare		14

MOCE-1	1992	N Pacific	C. Trees		1
MOCE-2	1993	N Pacific	-		6
MOCE-3	1994	N Pacific	-		7
NAPP Leg 3	1990	N Atlantic	G. Kraay and M. Veldhuis		2
NIOP-JGOFS B1	1992	Somali Basin, Gulf of Aden	-	19	2
NIOP-JGOFS B2	1993	Somali Basin, Gulf of Aden	-	-	3
OCT1999BRIAN	1999	Scotian Shelf	V. Stuart and E. Head		2
OLIPAC	1994	Equatorial Pacific	http://www.obs-vlfr.fr/cd_rom_dmtt/start_france_jgofs.htm	20	44
OMEX I VLD137	1993	NE Atlantic	http://www.bodc.ac.uk		7
OMEX I BG93-22 A	1993	NE Atlantic	-		1
OMEX I CD85	1994	NE Atlantic	-		4
OMEX I CD94	1995	NE Atlantic	-		2
OMEX I VLD154	1995	NE Atlantic	-		16
OMEX I DI216	1995	NE Atlantic	-		3
OMEX I DI217	1995	NE Atlantic	-		7
OMEX II CD105 B	1997	NE Atlantic	-		36
OMEX II CD110 B	1998	NE Atlantic	-		2
OMEX II PS237-1	1998	NE Atlantic	-		13
OMEX II CD114 A-B	1998	NE Atlantic	-		14
OMEX II BG9919 B-C	1998	NE Atlantic	-		15
POMME1	2001	N Atlantic	http://www.lodyc.jussieu.fr/POMME/		79
POMME2	2001	N Atlantic	-		85
POMME3	2001	N Atlantic	-		81
PROMOLEC	2000	Red Sea	J. Ras		10
PROSOPE	1999	NE Atlantic, Mediterranean Sea	http://www.obs-vlfr.fr/cd_rom_dmtt/start_france_jgofs.htm		62
ROAVERRS NBP 96-6	1997	Ross Sea	K. Arrigo and G.R. DiTullio	21	69
ROAVERRS NBP 97-9	1997	Ross Sea	-	-	53
ROAVERRS NBP 98-7	1998	Ross Sea	-	-	66
SAZ AU9706	1998	Southern Ocean (Subantarctic Zone)	S. Wright		27
SOMARE	2000	Tropical Atlantic	G. Kraay and M. Veldhuis		1
Transpacific Leg I (24°N)	1985	Trans N Pacific	R. Bidigare	22	39
Transpacific Leg II (47°N)	1985	Trans N Pacific	R. Bidigare	-	14
Vancouver	1996	Off Vancouver Island	V. Stuart and E. Head	23	2
WOCE P15S	1996	Trans S Pacific	G. R. DiTullio	24	33
					2419

A2: Origin of the data included in database #2

Project and Cruise	Year	Location	Source	Ref	n
ACSOE DI234	1998	NE Atlantic	http://www.bodc.ac.uk		39
AMT-2	1996	Trans Atlantic	G. Moore, S. Hooker, G. Westbrook	2	256
AMT-3	1996	Trans Atlantic	S. Hooker, J. Aiken, and R. Barlow	3	294
AMT-4	1997	Trans Atlantic	S. Hooker, J. Aiken, and L. Van Heukelem	2	296
AMT-5	1997	Trans Atlantic	S. Hooker, J. Aiken, and R. Barlow	2	302
AMT-6	1998	Trans Atlantic	-	3	263
AMT-7	1998	Trans Atlantic	S. Hooker, J. Aiken, and L. Dransfield	-	65
AMT-8	1999	Trans Atlantic	-	-	264
Bahamas Experiment C1	2000	Off Bahamas	S. Hooker, D. van der Linde, and C. Targa		37
Bahamas Experiment C4	2001	Off Bahamas	S. Hooker and L. Van Heukelem		82
BOFS STERNA DI198	1992	Bellingshausen Sea	R. Barlow / http://www.bodc.ac.uk		10
BOFS STERNA JCR2	1992	Bellingshausen Sea	-	7	23
GeP&CO A	1999	Trans N Atlantic - S Pacific	http://www.lodyc.jussieu.fr/gepco	25	117
GeP&CO B, C, D, E	2000	Trans N Atlantic - S Pacific	-	-	453
GeP&CO F, G, H, I	2001	Trans N Atlantic - S Pacific	-	-	492
GeP&CO J, K, L	2002	Trans N Atlantic - S Pacific	-	-	457
ICECOLORS 90	1990	Bellingshausen Sea	R. Bidigare		31
IRONEX I 93	1993	S Pacific	-	15	10
MLML 91	1991	Off Iceland coast	-		56
MOCE-1	1992	N Pacific	C. Trees		28
MOCE-3	1994	N Pacific	-		73
MOCE-4	1998	N Pacific	-		216
MOCE-6	2000	N Pacific	-		115
OMEX I CD84	1994	NE Atlantic	http://www.bodc.ac.uk		5
OMEX I CD94	1995	NE Atlantic	-		9
OMEX I VLD154	1995	NE Atlantic	-		56
OMEX I DI216	1995	NE Atlantic	-		15
OMEX I PS211	1995	NE Atlantic	-		62
OMEX II CD105 B	1997	NE Atlantic	-		20
OMEX II CD110 B	1998	NE Atlantic	-		41
OMEX II BG9815C	1998	NE Atlantic	-		20
OMEX II CD114A	1998	NE Atlantic	-		2
OMEX II BG9919B	1999	NE Atlantic	-		12
OMEX II BG9919C	1999	NE Atlantic	-		17
					4238

A3: Note infra-table

(1) Claustre, H., P. Kerherve, J. C. Marty, L. Prieur, C. Videau, and J. H. Hecq (1994), Phytoplankton dynamics associated with a geostrophic front: Ecological and biogeochemical implications, *J. Mar. Res., 52*(3-5), 711-742.

(2) Gibb, S. W., R. G. Barlow, D. G. Cummings, N. W. Rees, C. Trees, P. M. Holligan, and D. Suggett (2000), Surface phytoplankton pigment distributions in the Atlantic Ocean: An assessment of basin scale variability, *Prog. Oceanogr., 45*, 339-368.

(3) Barlow, R. G., J. Aiken, P. M. Holligan, D. G. Cummings, S. Maritorena, and S. B. Hooker (2002), Phytoplankton pigment and absorption characteristics along meridional transects in the Atlantic Ocean, *Deep Sea Res., Part I, 49*(4), 637-660.

(4) Cailliau, C., H. Claustre, and S. Giannino (1997), Chemotaxonomic analysis of phytoplankton distribution in the Indian sector of the Southern Ocean during late austral summer, *Oceanol. Acta, 20*(5), 721-732.

(5) Barlow, R. G., R. F. C. Mantoura, and D. G. Cummings (1999), Monsoonal influence on the distribution of phytoplankton pigments in the Arabian Sea, *Deep Sea Res., Part II, 46*(3-4), 677-699.

(6) Barlow, R. G., R. F. C. Mantoura, M. A. Gough, and T. W. Fileman (1993), Pigment signatures of the phytoplankton composition in the northeastern Atlantic during the 1990 spring bloom, *Deep Sea Res., Part II, 40*(1-2), 459-477.

(7) Barlow, R. G., R. F. C. Mantoura, and D. G. Cummings (1998), Phytoplankton pigment distributions and associated fluxes in the Bellingshausen Sea during the austral spring 1992, *J. Mar. Systems, 17*(1-4), 97-113.

(8) http://wwwold.nioz.nl/projects/dcm/

(9) Vidussi, F., J. C. Marty, and J. Chiaverini (2000), Phytoplankton pigment variations during the transition from spring bloom to oligotrophy in the northwestern Mediterranean Sea, *Deep Sea Res., Part I, 47*(3), 423-445.

(10) Bidigare, R. R., and M. E. Ondrusek (1996), Spatial and temporal variability of phytoplankton pigment distributions in the Central Equatorial Pacific Ocean, *Deep Sea Res., Part II, 43*(4-6), 809-833.

(11) Claustre, H., and J. C. Marty (1995), Specific phytoplankton biomasses and their relation to primary production in the Tropical North Atlantic, *Deep Sea Res., Part I, 42*(8), 1475-1493.

(12) Mackey, D. J., H. W. Higgins, M. D. Mackey, and D. Holdsworth (1998), Algal class abundances in the western equatorial Pacific: estimation from HPLC measurements of chloroplast pigments using CHEMTAX, *Deep Sea Res., Part I, 45*(9), 1441-1468.

(13) Mackey, D. J., J. Blanchot, H. W. Higgins, and J. Neveux (2002b), Phytoplankton abundances and community structure in the equatorial Pacific, *Deep Sea Res., Part II, 49*, 2561-2582.

(14) Higgins, H. W., and D. J. Mackey (2000), Algal class abundances estimated from chlorophyll and carotenoid pigments, in the western Equatorial Pacific under El Niño and non-El Niño conditions, *Deep Sea Res., Part I, 47*(8), 1461-1483.

(15) Special issue (1998), The Galapagos Iron Experiments: A tribute to John Martin, *Deep Sea Res., Part II, 45*(6).

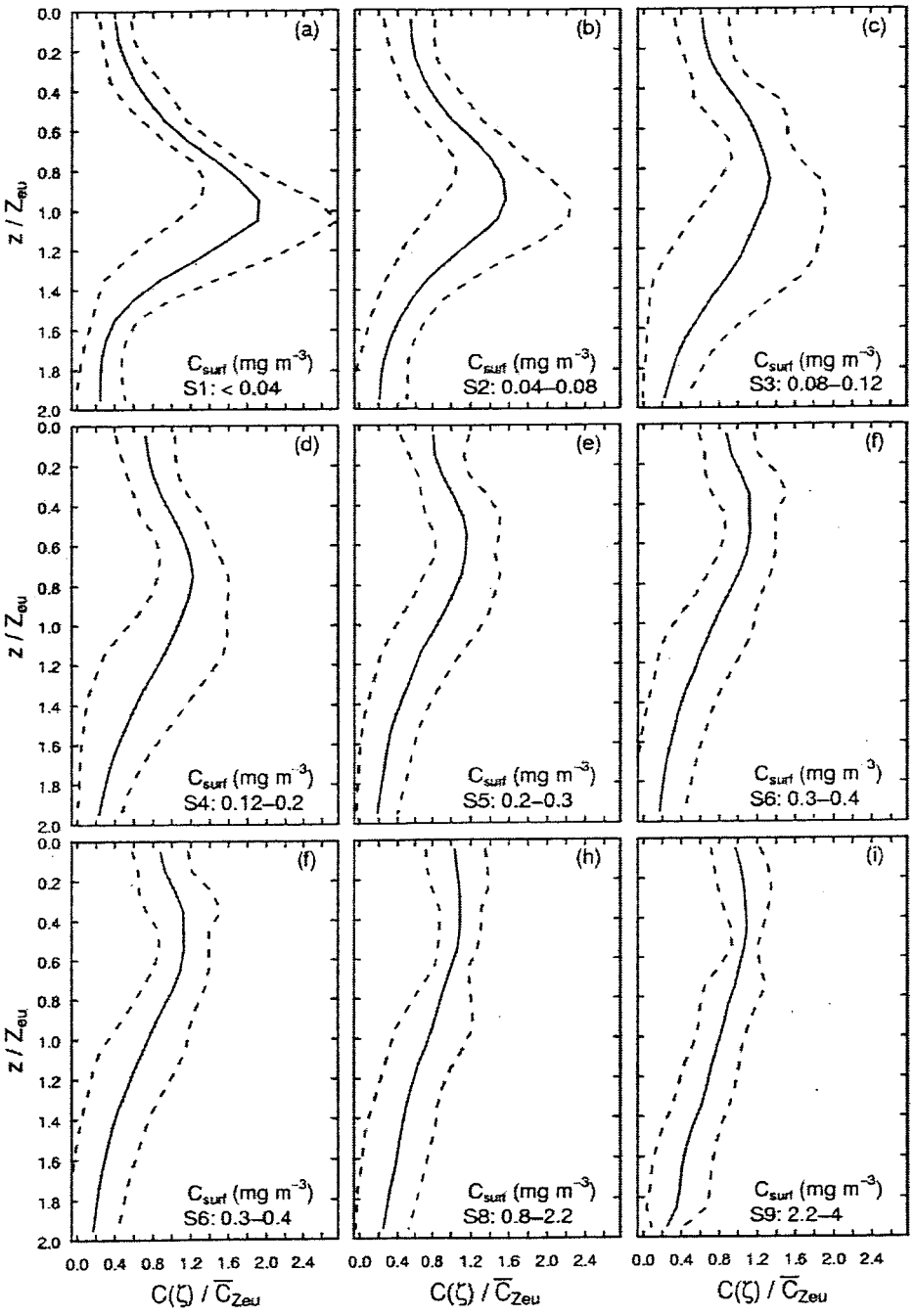
(16) Stuart, V., S. Sathyendranath, E. J. H. Head, T. Platt, B. D. Irwin, and H. Maass (2000), Bio-optical characteristics of diatom and prymnesiophyte populations in the Labrador Sea, *Mar. Ecol. Prog. Ser., 201*, 91-106.

(17) Vidussi, F., H. Claustre, B. B. Manca, A. Luchetta, and J. C. Marty (2001), Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern

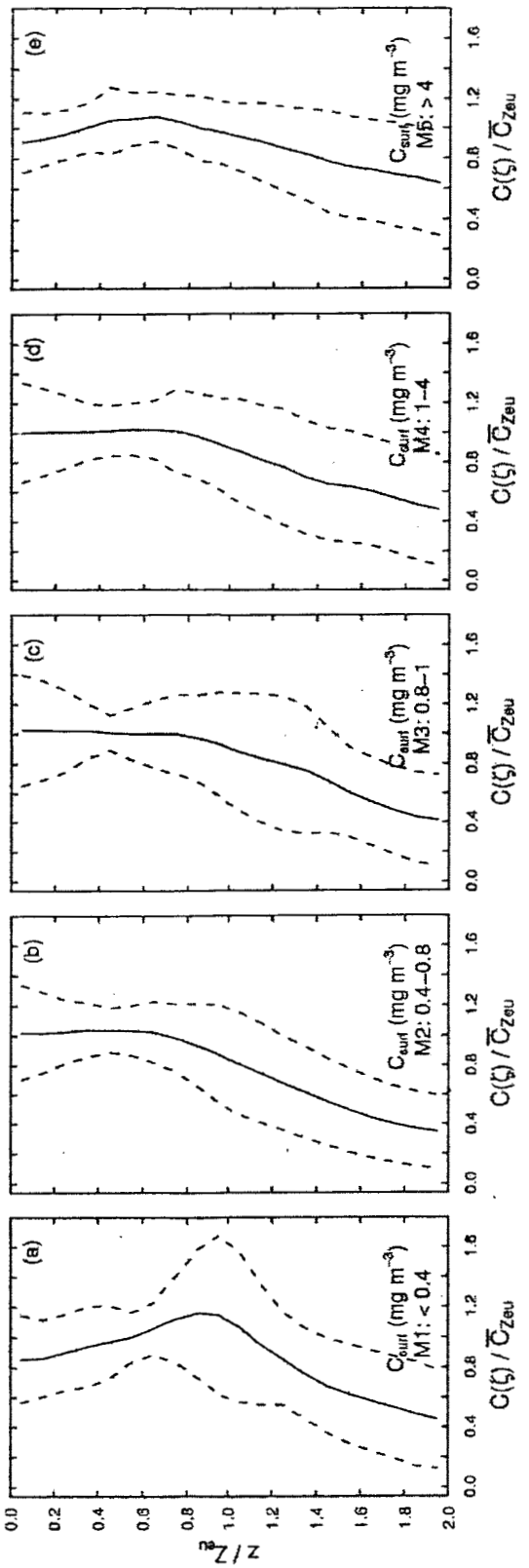
Mediterranean Sea during winter, *J. Geophys. Res., 106*(C9), 19,939-19,956.
(18) Stuart, V., O. Ulloa, S. Sathyendranath, T. Platt, G. Alarcon, H. Major, and E. J. H. Head (2004), Characteristics of phytoplankton populations in the upwelling system off the coast of Chile, *Rev. Chilena Hist. Nat.*, In press.
(19) http://wwwold.nioz.nl/en/facilities/dmg/niop/niop.htm
(20) Claustre, H., A. Morel, M. Babin, C. Cailliau, D. Marie, J. C. Marty, D. Tailliez, and D. Vaulot (1999), Variability in particle attenuation and chlorophyll fluorescence in the Tropical Pacific: Scales, patterns, and biogeochemical implications, *J. Geophys. Res., 104*(C2), 3401-3422.
(21) http://www.cofc.edu/~ditullio/rosssea.html
(22) Ondrusek, M. E., R. R. Bidigare, S. T. Sweet, D. A. Defreitas, and J. M. Brooks (1991), Distribution of phytoplankton pigments in the North Pacific Ocean in relation to physical and optical variability, *Deep Sea Res., 38*(2), 243-266.
(23) Stuart, V., S. Sathyendranath, T. Platt, H. Maass, and B. D. Irwin (1998), Pigments and species composition of natural phytoplankton populations: Effect on the absorption spectra, *J. Plankton Res., 20*, 187-217.
(24) DiTullio, G. R., M. E. Geesey, D. R. Jones, K. L. Daly, L. Campbell, and O. S. J. Walker (2003), Phytoplankton assemblage structure and primary productivity along 170°W in the South Pacific Ocean, *Mar. Ecol. Prog. Ser., 255*, 55-80.
(25) Dandonneau, Y., P. Y. Deschamps, J. M. Nicolas, H. Loisel, J. Blanchot, Y. Montel, F. Thieuleux, and G. Bécu (2004), Seasonal and interannual variability of ocean color and composition of phytoplankton communities in the North Atlantic, Equatorial Pacific and South Pacific, *Deep Sea Res., Part II*, In press.

Appendix B: Average dimensionless profiles of C, micro-C, nano-C and pico-C, and their standard deviations

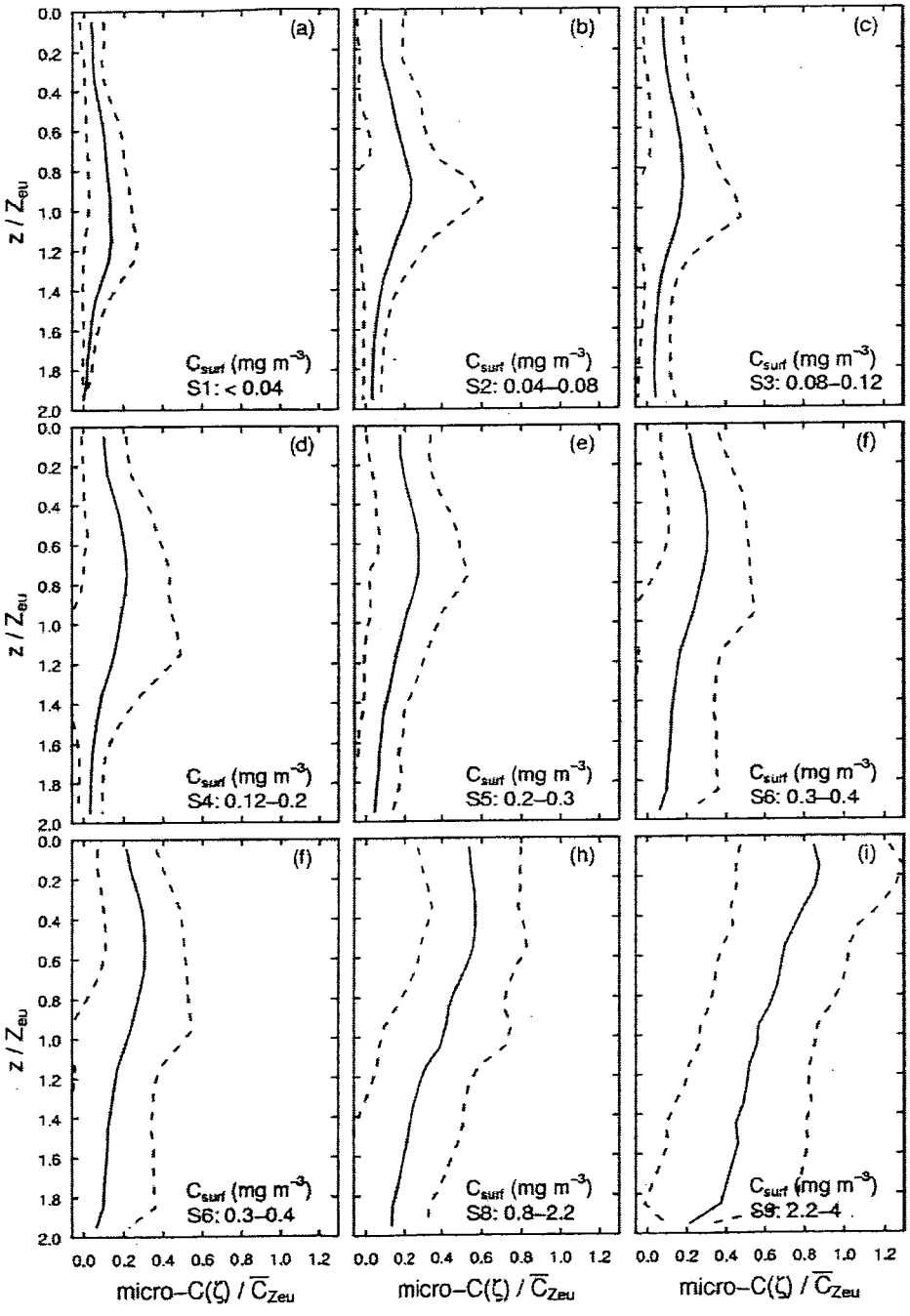
B1: Average dimensionless C profiles for each trophic category (S1 to S9) of stratified waters. The dashed lines represent the mean values ± sd.



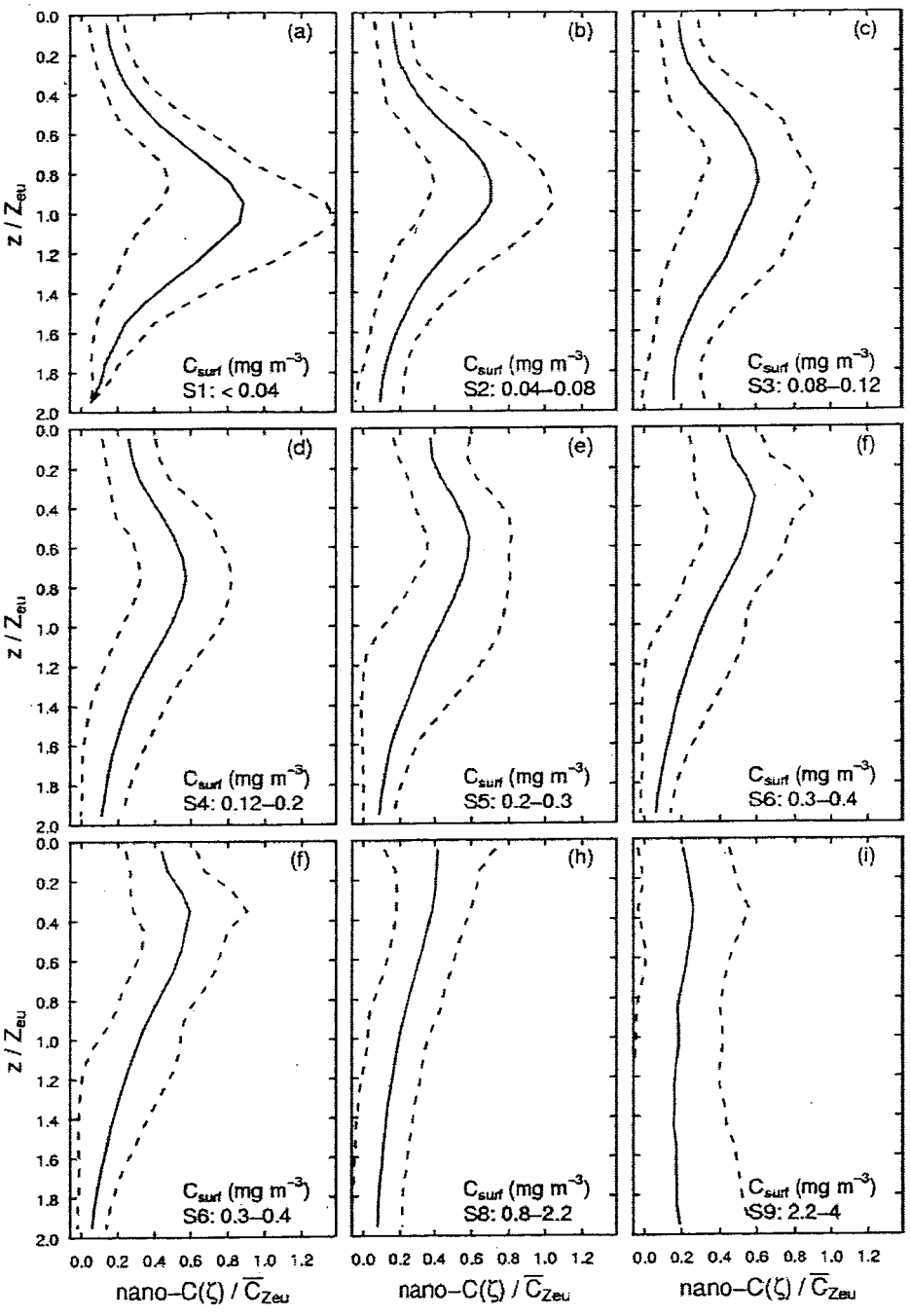
B2: Average dimensionless C profiles for each trophic category (M1 to M5) of mixed waters. The dashed lines represent the mean values ± sd.



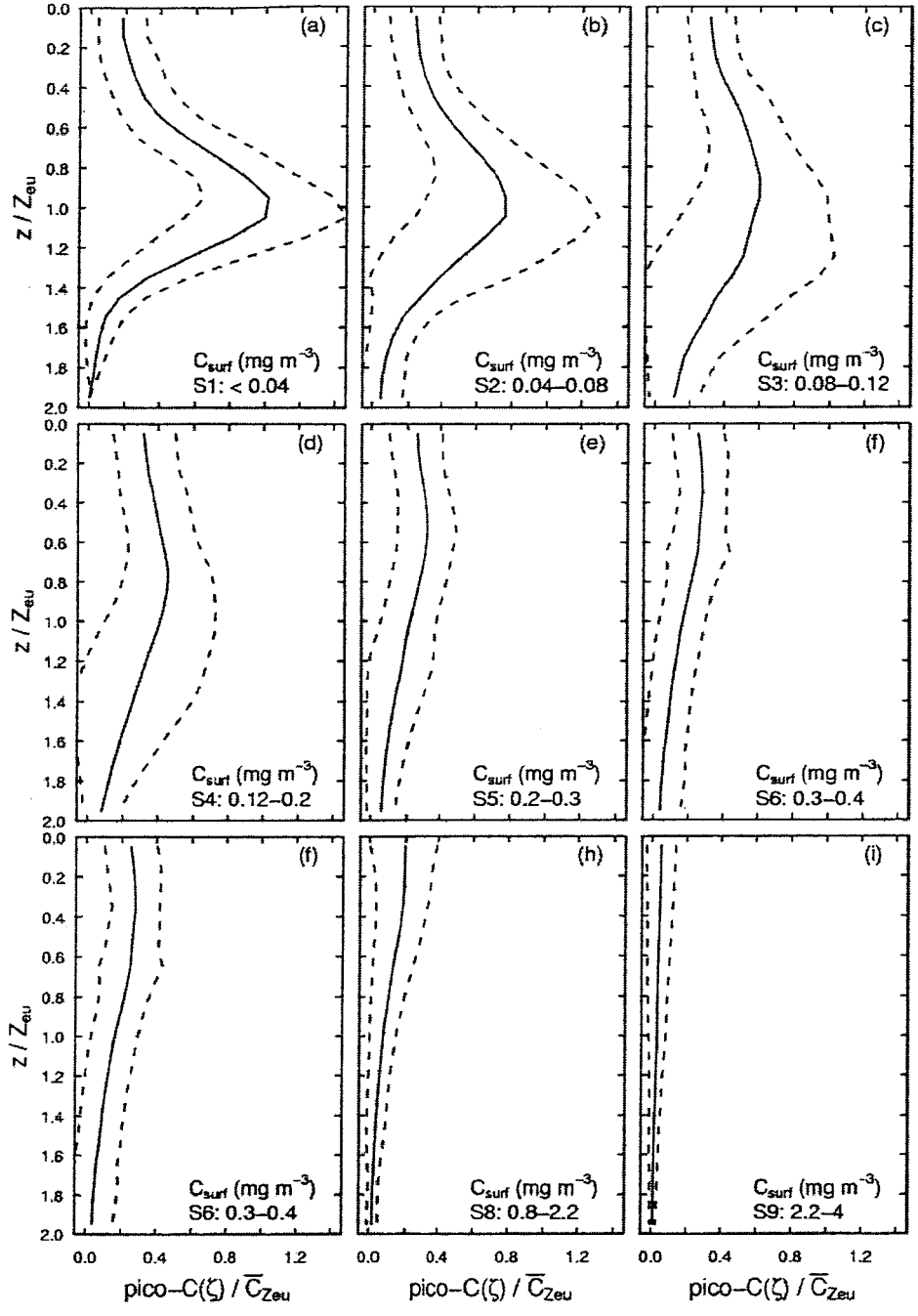
B3: Average dimensionless micro-C profiles for each trophic category (S1 to S9) of stratified waters. The dashed lines represent the mean values ± sd.



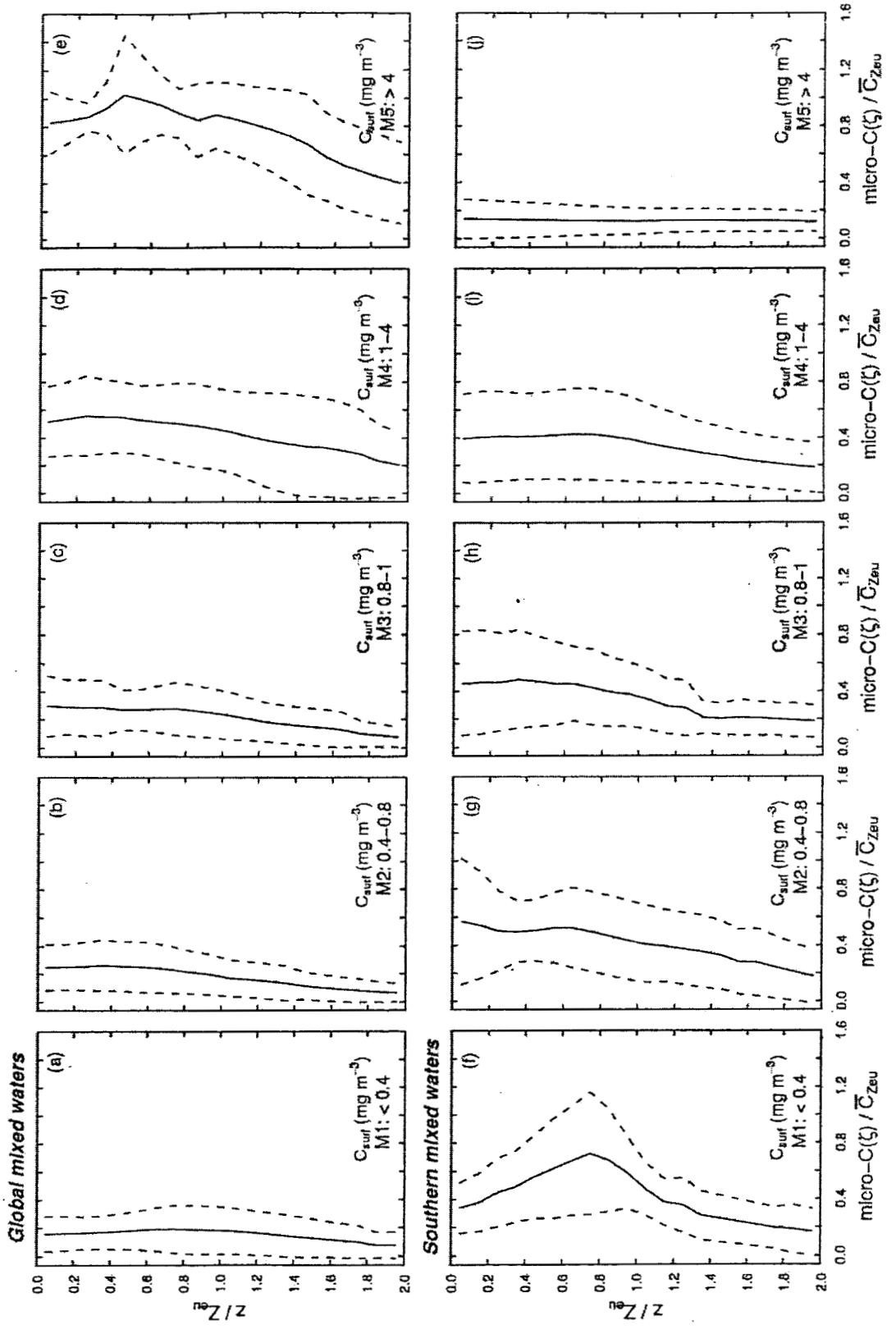
B4: Average dimensionless nano-C profiles for each trophic category (S1 to S9) of stratified waters. The dashed lines represent the mean values ± sd.



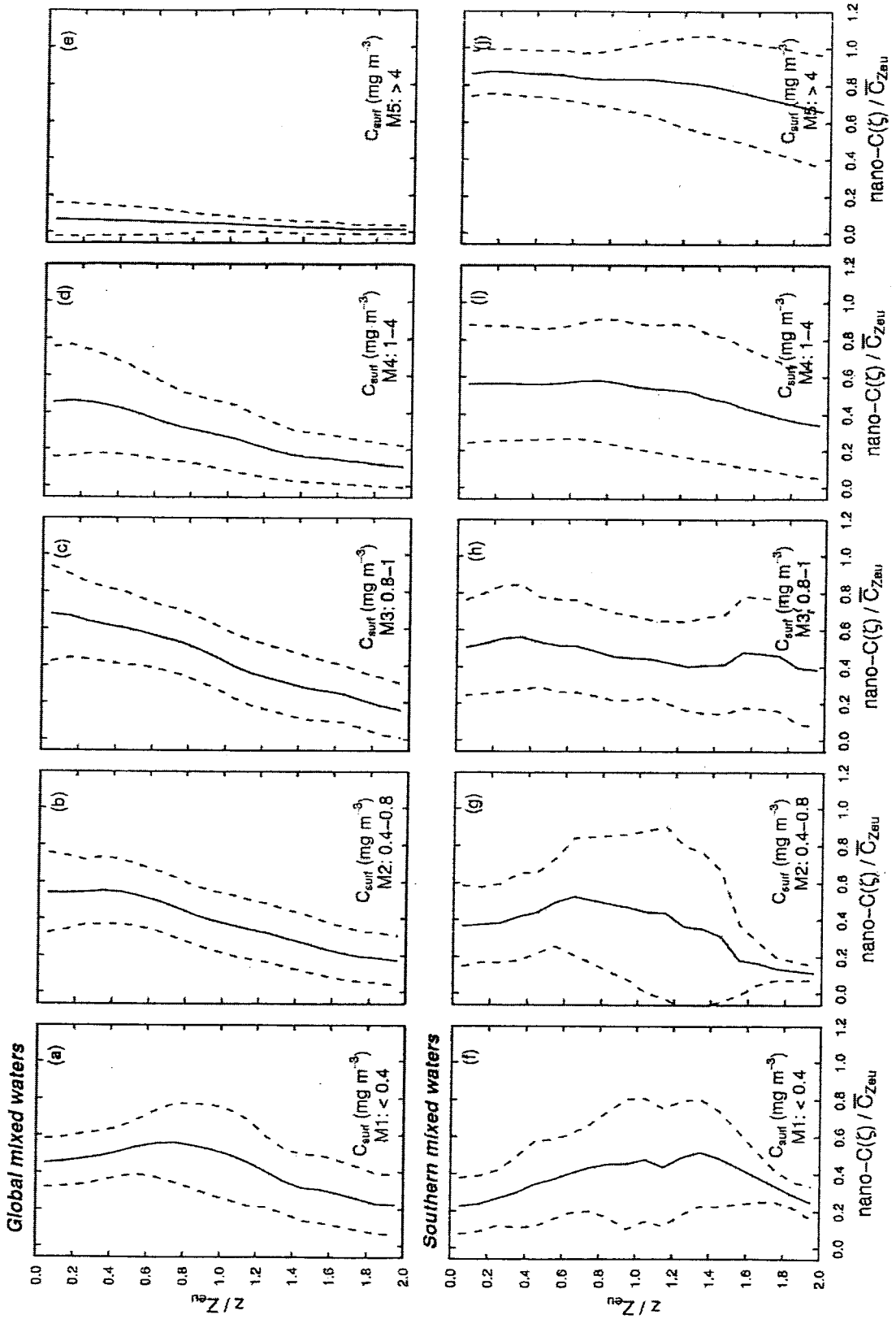
B5: Average dimensionless pico-C profiles for each trophic category (S1 to S9) of stratified waters. The dashed lines represent the mean values ± sd.



B6: Average dimensionless micro-C profiles for each trophic category (M1 to M5) of global (a-e) and southern (f-j) mixed waters. The dashed lines represent the mean values ± sd.



B7: Average dimensionless nano-C profiles for each trophic category (M1 to M5) of global (a-e) and southern (f-j) mixed waters. The dashed lines represent the mean values ± sd.



B8: Average dimensionless pico-C profiles for each trophic category (M1 to M5) of global (a-e) and southern (f-j) mixed waters. The dashed lines represent the mean values ± sd.